Research Report

The impact of sanitation marketing and the characteristics of Community Health Workers on sanitation marketing in Western Kenya

Introduction

Kenya has been struggling with the public health consequences of non-existent or improper sanitation systems for decades. This problem is alarming since inadequate sanitation infrastructure and insufficient hygiene habits lead to higher disease rates and mortality (Edward, Jung, Chhorvann, Ghee, & Chege, 2019). On the other hand, improved sanitation practices have been proven to prevent many diseases that impair working population and children (Ellis et al., 2020). In addition to the undoubtful immediate benefits of improved hygiene behavior, these benefits are expected to lead to long-lasting health improvements since the practices are passed from generation to generation (Edward et al., 2019). Yet, developing adequate hygiene habits needs to be complemented with sufficient sanitation infrastructure to work (Jenkins & Curtis, 2005). Personal hygiene habits are very closely interlinked with establishing latrine usage instead of the default open defecation (Jenkins & Curtis, 2005). Installing and using hygienic latrines is crucial as it lowers the exposure to fecal pathogens that cause health problems through malnutrition (Ellis et al., 2020). Moreover, oftentimes installing a latrine leads to the further development of hygiene habits and seeing a greater importance in good sanitation (Jenkins & Curtis, 2005). It is the steppingstone that is so needed for rural population to improve their overall hygiene habits and consequently improve their health.

Our research analyzes one of the widely-used practices to improve sanitation in rural populations – sanitation marketing. The research was conducted through a survey in two regions of Western Kenya – Busia and Bungoma, in which some of the community health workers self-identified as sanitation marketers involved in the program. Through an elaborate survey covering both regions, personal determinants of community health workers were collected alongside the results of sanitation-related outcomes of households and their involvement in community meetings. The paper uses descriptive statistics and regression analysis in order to answer the two main research questions: *Does sanitation marketing lead to improvement of sanitation outcomes?* and *To what extent do job and personal characteristics, including payment, motivation for the job and the status of a community health worker influence sanitation improvement of villagers?* In order to answer the research questions, we analyze the impact of sanitation marketing on sanitation improvements, namely improvements of sanitation systems and hygiene behavior as these are oftentimes connected, yet the outcomes are different, as well as CHW characteristics determining these improvements.

Because of the multifold benefits of good sanitation, the Kenyan government has been focusing on the promotion of hygiene habits through its policies. Among others, Kenya has launched its community health strategy in 2006 as a part of National Health Sector Strategic Plan (2005-2010 NHPP II). The most important role is played by community units (CU) that constitute for the primary response in the health system (Kawakatsu, Tanaka, Ogawa, Ogendo, & Honda, 2017). Each CU consists of a community health committee (CHC), two community health extension workers (CHEWs) and community health workers (CHWs) or the equivalent community health volunteers (CHVs) (Kawakatsu et al., 2017). CHWs are at the very core of each unit, being responsible for delivering health interventions to locals (Wangalwa et al., 2012). In fact, more countries in sub-Saharan Africa have adopted programs promoting community-based health following the successes of the pioneer program in Zimbabwe (Ballard et al., 2018). CHWs are mostly members of local communities, who are closer to them and understand their culture (Kawakatsu et al., 2015). This distinction and unique relation to the community might prove to be an important factor in succeeding when implementing health strategies compared to a third-party program implementation (Kawakatsu et al., 2015). Therefore, it is crucial to examine the role of CHWs in health and sanitation promotion alongside the individual determinants influencing their impact on sanitation promotion.

In line with the National Health Sector Strategic Plan, the Kenyan Ministry of Health, launched a Community Led Total Sanitation (CLTS) program on which several organizations developed their targeted programs including sanitation marketing training and promotion (UNU MERIT, Endline Report FINISH-INK Project, 2019). FINISH-INK with its partner organizations have developed a program building on the community health structures, employing mostly the existing CHWs, in order to improve the sanitation infrastructure and hygiene habits of the locals. The aim of the program was to create primary demand for sanitation in line with the CLTS program, without aiming at a specific type of sanitation system necessarily. The CHWs were trained to emphasize the durability of the sanitation system and the relevance of improved sanitation systems to community members. Similar sanitation marketing programs were implemented in many developing countries showing promising results in various forms of improved sanitation infrastructure and hygiene habits (Evans et al., 2014; Grier & Bryant, 2005; Jenkins & Scott, 2007). However, more empirical research is needed to establish a link between sanitation marketing and improved sanitation (Evans et al., 2014). Sanitation improvements, which are the outcomes of sanitation marketing programs, can be manifested in multiple ways. The improvements can comprise of installing a new sanitation system or improving an existing one, to usage of the sanitation system and its maintenance as a behavioral change. These are oftentimes intertwined, yet it is also important to study them separately to see, whether there are any significant differences between them. Based on the up-to-date research suggesting multifold sanitation improvements motivated by the sanitation marketing implementation, this paper poses the first research question, *Does sanitation marketing lead to improvement of sanitation outcomes?*

Existing academic literature has mostly focused on the role of community health workers and their efficiency in implementing desirable programs as will be presented in the literature review section. However, very few have emphasized the role of their task-specific motivation, receiving payment or their status in connection to actual sanitation outcomes at the household level. Yet, this paper supposes that personal motivation of a local community member who has a significant authority – CHW or CHV might have an inspiring impact on local communities (Kawakatsu, Sugishita, Kioko, Ishimura, & Honda, 2012), mainly through community meetings – barazas. We also study other characteristics of a CHW which might influence their job effectiveness, such as the payment they receive and whether they perceive being a CHW marketer as their main job which is related to their status. Henceforth, this paper poses its second main research question*: To what extent do job and personal characteristics, including payment, motivation for the job and the status of a CHW influence sanitation improvement of villagers?*

In the beginning, the paper reviews the relevant academic literature while providing a useful theoretical background for the research. Later, it presents detailed methodology, including survey and program design. Furthermore, we present variable creation of the variables relevant for our research together with descriptive statistics that provide a basic insight into the data. Consequently, this paper provides a short overview of the analysis and inferential findings in the Results section. The main finding is the positive effect of sanitation marketing on both sanitation outcomes. However, our study found that none of the CHW characteristics were significant in determining sanitation outcomes. Furthermore, we discuss the implications of these findings and potential limitations of the research. Lastly, the paper draws a conclusion on our results.

Literature review and Theoretical Framework

Several academic studies have examined the efficiency of community-led health programs and social marketing in particular. Social marketing can be defined as "the use of marketing to design and implement programs to promote socially beneficial behavior change" (Grier & Bryant, 2005). Thus, sanitation marketing can also be considered a form of social marketing as it seeks to improve sanitation

behavior and infrastructure. Overall, the use of social marketing as a practice in improving health in developing countries has been increasing (Evans et al., 2014). Through marketing hygiene behaviors or health-related products such as sanitation systems, social marketing programs are aiming to improve health of communities (Grier & Bryant, 2005), while sharing crucial aspects with commercial marketing (Evans et al., 2014). Both commercial and social marketing seek to exchange value between a marketer and a consumer, which forms the basis for a successful marketing program (Grier & Bryant, 2005). In the case of social marketing in public health, the value is associated with a particular health behavior or facility installment (Grier & Bryant, 2005). Applied to the FINISH INK´s sanitation marketing project, the main product is a new or an improved sanitation system that households can adopt. However, secondary products are equally important – the improvement of sanitation behavior is also considered a valuable exchange that can be promoted by sanitation marketer. Thus, our study examines both outcomes aiming to compare them and see whether sanitation marketing has a different effect on each.

None of marketers´ efforts can be seen in isolation from the social nature of social marketing. Most importantly, social marketing is a community-led strategy, concentrating on social change through community, not just mere communication and promotion (Jenkins & Scott, 2007). When motivating change, it is shown that community-led strategy is one of the most effective practices in implementing health-improving interventions (Wangalwa et al., 2012). However, the results of behavioral change following social marketing are mixed (Evans et al., 2014). Studies employing social marketing have mostly found that marketing has created a higher level of awareness of health threats and solutions such as the marketed products, but this does not necessarily lead to a change in behavior (Evans et al., 2014). One of the reasons for the mixed results is that improving hygiene habits is not simply a matter of a narrow marketing campaign, but a set of more factors that must tailor to the needs of locals, placing community at its core (Grier & Bryant, 2005). The factors can range from what is seen as the main health product distinguishing between behavior change and tangible facilities, the place and time during which social marketing is implemented, including informal community interactions, to price of a given product (Grier & Bryant, 2005). All of these play an integral part in determining the success of a social marketing program.

Moreover, one of the crucial factors determining the success of a community-led strategy, such as sanitation marketing, is who from the community initiates the program and takes the responsibility for its implementation (Y. Kawakatsu, Sugishita, Kioko, Ishimura, & Honda, 2012). It is oftentimes the local CHW, who ensures that the social marketing program tailors well enough to the needs of the local community (Grier & Bryant, 2005). Yet, studies analyzing the impact of CHWs on the degree of various health improvements of locals show mixed results (Edward et al., 2019; Kawakatsu et al., 2017; Olayo, Wafula, Aseyo, Loum, & Kaseje, 2014). Because of the mixed results, it is not only important to study the efficiency of social marketing programs, but also what determines the outcome at the CHW level. CHWs are oftentimes the only direct service providers in communities (Kawakatsu et al., 2017) and they are also involved in local communities informally (Aseyo et al., 2018). Therefore, their impact could be said to be two-fold: they provide service directly via door-to-door marketing or community health meetings, but they also formally and informally engage with locals, exchanging valuable information motivating change (Kawakatsu et al., 2012). Thus, the overall performance of CHWs is influenced by several factors.

A study has found that older CHWs tend to perform better than younger ones, suggesting a link between the community status of a CHW in a village and their ability to motivate change (Kawakatsu et al., 2012). Yet, it is also possible that the position alone can lead to a greater status in the community, if being CHW is their main job (Alam, Tasneem, & Oliveras, 2012; Y. Kawakatsu et al.,

2012). This could be especially relevant for inducing change socially, as locals report that an increased status that comes with improving one´s sanitation system is one of most important motivators for its adoption (Jenkins & Curtis, 2005). Subsequently, CHW´s greater authority and status in a local community can help motivate locals to change their sanitation systems and behaviors more effectively.

Another factor impacting the performance of a CHW is their motivation. The study of Aseyo and colleagues that was focusing on unpaid CHWs reported that limited time and material resources lead to decreased CHW motivation, which is one of the main reasons for their poor performance (Aseyo et al., 2018). Time spent on performing direct social marketing activities is a good indicator of CHW´s motivation for the job and also a good determinant for the success of sanitation improvement of the household (Aseyo et al., 2018; Ashraf, Bandiera, Davenport, & Lee, 2020). Generally speaking, more motivated CHW marketers tend to spend more time on sanitation marketing and engage with their work more than when they lack the motivation (Ashraf et al., 2020). Motivation for the marketing job energizes the direct action of a CHW when performing their job, which is absolutely crucial for inspiring change in the local community (Aseyo et al., 2018). Thus, it is expected to be closely related to sanitation outcomes of the community where a motivated CHW operates.

Overall, unpaid CHWs reported low levels of motivation due to no payment and the need to work an additional job to make living (Aseyo et al., 2018). Not only does this influence the time spent on marketing, but it might cause unnecessary existential stress to the CHW and exacerbate their performance (Aseyo et al., 2018). It is thus important to examine the relationship between payment and CHW performance. If the wage of a CHW is insufficient, their motivation and efficiency decreases (Aseyo et al., 2018). A CHW is essentially an employee and motivating salary increases motivation and attracts talented individuals who prove to be more efficient in their jobs (Ashraf et al., 2020). Additionally, the job characteristics of CHW´s work, such as having a positive impact on a local community, have a strong potential to attract prosocial individuals with high intrinsic motivation, which in turn leads to their higher performance as well (Ashraf et al., 2020). Yet, the motivation and the will to perform one´s job suffer if the payment is too low or even non-existent (Aseyo et al., 2018).

Lastly, CHW marketers can also influence their peers through community activities. Such organized meetings aiming to engage the local communities with new sanitation programs oftentimes take form of barazas. A baraza is a public meeting, which is led by a local authority, who is oftentimes the local CHW (Naanyu et al., 2011). Several studies have analyzed the role of health barazas in community-based health intervention programs finding a positive effect on health behavior (Njomo et al., 2017; Ojakaa et al., 2011). Barazas are one of the means for CHWs to establish social rapport with locals while communicating important ideas (Ojakaa et al., 2011), which has been found to be one of the most important factors determining CHW´s performance (Ashraf et al., 2020). By going to regular meetings, locals are more exposed to community pressures regarding sanitation change (Jenkins & Curtis, 2005). Moreover, the meetings can also elicit disgust that might lead to higher want to adopt new sanitation systems such as latrines (Jenkins & Curtis, 2005). What is crucial for sustaining and organizing these informative meetings is the motivation of the local CHW authority (Njomo et al., 2017). Here comes the second level of a CHW´s work – the social impact that should mediate behavioral change especially. Yet, this also ties into how CHW perceived his role, the CHW needs to show enthusiasm and motivation to organize these meetings (Kawakatsu et al., 2017; Njomo et al., 2017).

This research contributes to the existing academic literature by analyzing the individual determinants of CHW on their performance, which can be summarized as the status associated with the job, the motivation for the job demonstrated through time spent on it and the satisfaction of CHW´s monetary needs. *What is believed to have an especially high impact on the villagers´*

sanitation outcome is the high motivation of a CHW leading to organization of barazas. The research might serve as a useful guide for policy recommendations regarding motivating CHW and potentially leading health change through community meetings as well; if their effect is proven significant.

Methodology

In this section, we provide the description of the sanitation marketing program, from which the data were collected. Following the detailed presentation of the program design, the data collection method is described. Moreover, we pose hypotheses this research tries to analyze throughout. In order to examine the hypotheses using the collected data, we present basic data manipulation and preparation for further analysis. Additionally, the statistical procedures used when performing the analysis are described in the end of the Methodology section.

Intervention

This research uses data collected in a cross-sectional observational study following a sanitation marketing program in Western Kenya to answer the two research questions, *Does sanitation marketing lead to improvement of sanitation outcomes?* and *To what extent do job and personal characteristics, including payment, motivation for the job and the status of a community health worker influence sanitation improvement of villagers*. The study took place in 80 villages in the counties Busia and Bongoma, each county accounting for half of the villages. Following the Kenyan government´s CLTS initiative, FINISH INK and partner organizations established a sanitation marketing program taking place in 2015-2017, in which they trained locals in the Busia county, Teso North constituency, to promote installment of a new sanitation system such as a latrine. This area and households located in this area are referred to as the treatment group. Not only were the villagers trained to market the importance of having a latrine, but specifically the importance of having an improved, good and durable system, which was the main message of the project-specific marketing.[1] The villagers trained for this role were mostly local CHWs, however some of them were village elders or local farmers. These sanitation marketers were then directly approaching households on a door-to-door basis, promoting the installment of a better sanitation system. Sanitation marketing was performed as a job, with each sanitation marketer being primarily responsible for their location, yet being able to travel to a different location. After the program being in place for a year, results were collected by Amref (FINISH INK´s partner organization) interviewers via multiple surveys. The interviewers were chosen for their location randomly.

On the other hand, FINISH INK´s sanitation marketing program had not taken place at the time of collecting answers in Bungoma constituencies Sirisia and Bumula. Thus, they are jointly referred to as a no treatment or a control group. This research paper also analyzes the data collected from a random sample of villages in these locations to serve as comparison to the villages located in Busia that participated in the sanitation marketing program initiated by FINISH INK and partner organizations.

Data Collection – Questionnaire

Overall, villages located in both treatment and no treatment area were randomly selected for evaluation. The data was collected through several surveys, out of which 3 have crucial relevance for this research. The Household survey collected data on the household level concerning hygiene

---

[1] It is possible that spillover effects and activities initiated by other NGOs could have happened in both areas as some of the self-identified marketers in Bungoma area reported that they received training from a different organization. However, Amref´s sanitation marketing program was distinguished by training marketers to emphasize the importance of durability and the relevance of improved sanitation systems with more background information

behavior, sanitation installment and belief in sanitation importance among villagers. Moreover, the Village survey was also conducted in all villages, however it was focusing on the village-level interventions, interviewing the local CHW. Thirdly, the CHW survey was used to interview local self-identified sanitation marketers asking questions related to their marketing job[2]. Altogether, the data was collected from 1751 households in these counties as well as 122 sanitation marketers and 80 villages. Out of the 122 sanitation marketers, 80 also act as local CHWs, who this research focuses on. The study discounts sanitation marketers who are not simultaneously CHWs with the aim to avoid unwanted variations in results. This section presents the main data analysis strategies the research used as well as detailed steps of analyzing the data in order to answer the research question via evaluating hypotheses and drawing statistical inferences.

Each questionnaire for data collection was focusing on a different level of measurement. The Household questionnaire, reported in the Household dataset, interviewed randomly chosen households in the sample of 80 random villages. The questions in the survey were answered only by one member of the household and they were mostly concerned with sanitation outcomes on an individual household level, such as installing a new system or not. However, some of them were answered directly by the interviewer, asking them for an independent observation as directly indicated in the question. These questions mostly serve to evaluate success of sanitation marketing and CHW performance, as they capture the sanitation outcomes.

The information concerning CHWs in both constituencies was collected through 2 questionnaires. One questionnaire was answered by self-identified sanitation marketers, who could also self-identify as CHWs, and the data is reported in the dataset CHW. This questionnaire was focused on their subjective view of their performance, their payment, and motivation among other factors. The second questionnaire relevant for CHW characteristics was the Village questionnaire, which contained a broad range of questions as reported in the Village dataset. This questionnaire was answered by multiple people in the village, including the village elder and the local CHW. The questions were concerned with village-wide practices such as baraza meetings. Importantly, each questionnaire asked its respondents to report their village number, allowing us to pair households to villages operated by the corresponding CHWs.

Hypotheses

Following the basic layout of the research, we define 6 hypotheses:

Hypothesis 1a: The households that received sanitation marketing show better hygiene behavior.

Hypothesis 1b: The households that received sanitation marketing were more likely to change their sanitation systems to more advanced ones.

Hypothesis 2a: A frequent organization of health barazas, initiated by a motivated CHW, leads to better hygiene behavior and to more positive change of household sanitation facilities since barazas act also as a communication channel.

Hypothesis 2b: The satisfaction of CHW´s budgetary constraints through the sanitation marketing job leads to better hygiene behavior and to more change of household sanitation facilities.

Hypothesis 2c: More CHW´s motivation for their job leads to better hygiene behavior and to more positive change of household sanitation facilities.

---

[2] Locals who identified their activity as "sanitation marketing" participated in the survey. There was no formal requirement for what activity constitutes for sanitation marketing and whether or not, and by whom, they had to be trained

Hypothesis 2d: The status associated with seeing CHW sanitation marketing as their main job leads to better hygiene behavior and more change of household sanitation facilities.

Data Preparation

To analyze the data, all the appropriate datasets are combined by matching relevant information based on village ID numbers. The data used in this study was collected in three different questionnaires, therefore it was initially coded in three different datasets: Household dataset concerned with household level answered mainly by locals, Village dataset concerned with village-wide level mainly answered by village elders and CHWs, and Marketer dataset concerned with sanitation marketing answered by self-identified marketers who are oftentimes also CHWs. Therefore, first, we need to combine all the relevant datasets. Each marketer was responsible for one village, therefore based on the village ID, we can identify and pair the marketers with households in that village and the answers from the CHW in that village, who is the same person. We checked that there is at least one CHW marketer per village, which enables us to use village ID as an indicator for matching. Those villages or marketers who did not have a village ID that could be paired with other questionnaires were discounted, which decreased the number of household-level observations from 1751 to 1556. Moreover, not all marketers identified themselves as CHWs, and since the second aim of this research is to analyze the relationship between the individual characteristics of CHWs and their performance, these observations were not used in the data analysis. It was 8 marketers altogether who did not identify themselves as CHWs that were not used in the merged dataset used for analysis.

The pairing of datasets holds several important assumptions allowing us the merge of data. Firstly, we assume that the households located in a village with a certain village ID received sanitation marketing from their local CHW marketer, who reported the same village ID. Thus, the village ID is a correct matching indicator to measure the impact of a CHW on the area that he/she really worked in. Furthermore, for the sake of being able to reasonably divide and analyze the data, we assume that all of the households that belong to a village, which received sanitation marketing treatment, did indeed receive treatment. And all the villages that are in a region that received treatment, received treatment. Lastly, to be able to compare the results, we need to assume that there are no significant differences between Bumula and Sirisia, both part of non-treatment group.

Data Analysis – Statistical Procedures Used

Aiming to answer the research questions proposed and test the hypotheses, this research uses both descriptive and inferential statistics. Firstly, descriptive statistics is employed to define variables properly and study their overall character and distributions. This paper uses mostly Chi-squared statistics to compare categorical outcomes of categorical variables as belonging to a treatment or non-treatment group. Similarly, Wilcoxon rank sum test with continuity correction is used for descriptive comparisons of quantitative variables.

Secondly, regression models identifying the predictive nature of independent variable are used to test their effect. This paper provides 2 groups of regression models to answer the two research questions and testing hypotheses posed, divided by the outcome they analyze. First, there are two regression models that test the effect of receiving treatment or not receiving treatment on sanitation outcomes: regression 1 analyzes the effect of sanitation marketing on hygiene behavior using linear regression; and regression I analyzes this effect on the change of a sanitation system using probit binary regression. Furthermore, each of the two regression models is further refined as the number of barazas and CHW characteristics are added into each model. The regression models are grouped based on the sanitation outcome they analyze, yet the predictor variables are the same for both.

Variable Creation and Descriptive Statistics

In order to accurately describe data and test hypotheses proposed, we need to define variables in a greater detail. This section illustrates variable creation taking into account unreported data in the dataset to provide precise results. Moreover, we also descriptively analyze each variable, identifying the missing values, differences in treatment and control groups while testing for their significance as a pre-requisite for further regression analysis.

To describe data accurately and to be able to compare the two main groups of this research, we divide observations based on whether they received sanitation marketing (treatment group) or not (control group). Firstly, observations belonging to different constituencies were recoded under the assumption that all households located in Teso North constituency received treatment in the form of sanitation marketing, and households located in Sirisia and Bumula constituencies did not receive sanitation marketing treatment from the FINISH INK´s partner organizations. Observations in the *Area* variable were therefore respectively recoded into either "treatment" or "no_treatment" instead of real constituency names. This variable serves as an independent variable for research question I and II. The split is also referred to as *Area* throughout the paper, based on where the households are located. This variable is further used to compare the effects and distribution of other variables.

Dependent Variables

Hygiene Behavior

The first dependent variable hygiene behavior (coded as house_lat_comb) is created, representing the sanitation-related hygiene behavior outcome. It captures the usage and the maintenance of sanitation systems as they were reported by the interviewer evaluating them on the scale from 0-2 representing:

0   No
1   Somewhat
2   Very much so

This variable was created by taking means of two questions reporting the two important components of hygiene behavior – the usage of the system and how clean and well-maintained it appeared as reflected in the original questions: "*Are there signs of usage (such as a footpath from the house to the latrine)?*" and "*Does the latrine seem clean and well-maintained?*". The first question represents whether the latrine looks that it is used, which is an important information adding a more realistic approach to the overall sanitation improvement rather than just reporting number of facilities installed or sold. The second question captures the maintenance dimension. If locals really care about sanitation improvement, they are expected to maintain the latrine in a good state (Jenkins & Curtis, 2005). Since both of these variables are concerned with apparent usage reported by the interviewer, they were further coded into a single combined variable *hygiene behavior* by taking means of both questions for treatment and no treatment group. More details on combining these two variables are available in the Appendix.

Hygiene behavior was reported by households located in both treatment and no treatment areas. As shown by the variable representing hygiene behavior of a household, it is visible in Table 1 that the treatment group of households had a slightly higher mean than the no treatment group, although median was the same. There were around 80 missing observations in both groups. We wanted to see whether the differences between groups were significant. Thus, after testing for the normality of both distributions with Shapiro-Wilk normality test, we conducted the Wilcoxon rank sum test as both samples were not normally distributed. The Wilcoxon rank sum test with continuity correction had a significant result, showing higher mean of sanitation systems usage in households

that received treatment (n = 1556, p < 0.001). Therefore, the descriptive results suggest that receiving sanitation marketing treatment could potentially increase hygiene behavior.

Table 1 - Hygiene Behavior

	no_treatment	**treatment**
Mean	1.459	1.569
Median	1.5	1.5
SD	0.413	0.424
n (Missings)	756 (81)	800 (78)

Change of sanitation systems

Moreover, the dependent variable capturing the change in sanitation system of a household from its default version was created. This variable was coded as house_lat_new and is referred to as *Change of a sanitation system*, standing for installment or improvement of sanitation facilities as opposed to no action. It represents the locals´ answers to the question "Have you newly built or improved this latrine since 2015?". Since it was locals who answered the question, this variable is more subjective than the previous variable *hygiene behavior* that was reported by an interviewer. The answers for the *Change of a sanitation system* are represented on binary levels: where 0 stands for no change, which means that the household kept its the sanitation system intact. Level 1 was assigned to any change, including improving an existing system, installing a new one or installing a new one and further improving it. Originally, the variable had 4 possible levels, which can be found in the Appendix with the respective coding and prevalence in the sample as well as details on the use of inconclusive answers.

As depicted in Table 2, there are 83 and 82 missing answers in the respective groups, which were not used for further analysis of differences between the two samples. Out of the no treatment area group, 361 households changed their sanitation system to a better one, which is 54%. In the sanitation marketing treatment group, 445 households changed their sanitation system, representing 62% of the households in the group.

In order to test if the differences between households in regions that received sanitation marketing treatment and those who did not are significant, we performed the non-parametric Pearson´s Chi-squared test with Yates´ continuity correction. The test showed a significant difference in sanitation system change for households that received sanitation marketing and those who did not receive it (n= 1556, p = .002). Areas that did not receive sanitation marketing had a higher number of keeping their old system intact as opposed to those who received treatment.[3]

Table 2 - Change of Sanitation System

	No treatment	Treatment
0	312 (46%)	273 (38%)
1	361 (54%)	445 (62%)
Observations (Missings)	756 (83)	800 (82)

---

[3] Chi-square residuals, which are the differences in installment or non-installment that differ between treatment and non-treatment groups can be found in the Appendix

Independent variables

Data concerning CHWs and frequency of barazas can be viewed as independent variables that are further used in the regression analysis. In order to create relevant variables several steps were taken. Building on creation of the variables, basic descriptive statistics were performed for greater insight into the variables that can be incorporated into the regression analysis in next sections.

Barazas

Firstly, independent variable *number of barazas* (barazas_nr.x) is created. This quantitative variable reflects numeric answers to the question concerning local CHW, "In the last year, how many such meetings (health barazas) took place?". This question was answered by the local CHW reporting the number of barazas[4].

Table 3 presents the basic descriptive statistics of health barazas overall as well as reported by group based on receiving or non-receiving sanitation marketing treatment. Both the median and mean were higher in the no-treatment area than in the treatment area. After testing for the normality of both distributions with Shapiro-Wilk normality test, we conducted the Wilcoxon rank sum test as both of the samples were not normally distributed. The Wilcoxon rank sum test with continuity correction had a medium-significant result, showing higher mean of health barazas in the areas that did not receive sanitation marketing treatment than those who did (n = 1556, p = 0.02).

Table 3 - Number of Barazas

	no_treatment	treatment
Mean	3.59	3.39
Median	3	2
SD	2.02	2.06
n (Missings)	756 (0)	800 (0)

Time spent on sanitation marketing

Secondly, independent variable capturing the percentage of a CHW´s time spent on sanitation marketing is created. Self-identified CHW marketers answered to the survey question, "How much of your time do you spend on sanitation marketing? Give %." The answers were coded as:

0	0%
1	10%
2	20%
3	30% etc.

As another independent variable, *time spent on sanitation marketing* (coded as time_san) was analyzed as reported by CHW marketers operating in both areas, that received Amref sanitation marketing treatment and those who did not. Table 4 presents the basic descriptive statistics of time spent on sanitation marketing overall as well as reported by group based on receiving or non-receiving sanitation marketing treatment. The mean was slightly higher in the no-treatment area than the treatment area, with median being the same. However, the SD was also a bit higher in the no-treatment area compared to the treatment area. Moreover, after testing for the normality of both distributions with Shapiro-Wilk normality test, we conducted the Wilcoxon rank sum test as both of

---

[4] The question was initially not directed at health barazas specifically, yet it was included among questions inquiring about health practices and the topic of health. Also, health was the single most commonly regarded issue discussed at these meetings in further questions.

the samples were not normally distributed. The Wilcoxon rank sum test with continuity correction had a significant result, showing higher mean of CHWs in the areas that did not receive sanitation marketing treatment compared to those who did (n = 1556, p-value < 0.001).

Table 4 - Time spent on Sanitation Marketing

	No_treatment	Treatment
Mean	4.24	3.78
Median	4	4
SD	1.68	1.5
n (Missings)	756 (0)	800 (0)

Main job

Additionally, we are interested if the CHWs perceive sanitation marketing as their main job, which is an independent variable reflecting on the associated status of a CHW marketer. The answers were coded as: 0 representing that the marketer does not see sanitation marketing as their main job and 1 representing that the marketer considers it their main job at least partially.

The variable *main job* (no_job_san) was created as an answer to a question, "Do you see sanitation marketing as your main job?". Originally, the answers were in form of: 1 – Yes, 2 – It is part of my main job, 3 – No, it is a secondary activity. However, the answers 1 and 2 were combined as they reflect that the CHW marketer sees marketing as their main job at least partially. Thus, in a further analysis the answers reporting "No, it is a secondary activity" are coded as level 0 of the binary variable *main job* and 1 and 2 are both coded as level 1.

When comparing if the CHW marketers see sanitation marketing as their main job based on whether they operate in the area that received sanitation marketing treatment compared to no treatment, we compare the distribution of the answers. There were no missing values and the percentages in both groups were similar, indicating that sanitation marketers mostly view sanitation marketing as a part of their main job, as shown in Table 5, where 72% of those who operate in the no-reatment areas see sanitation marketing as their main job as opposed to 84% of those working in treatment areas. The relation between seeing sanitation marketing as their main job and receiving sanitation marketing treatment or not was significant as tested by Chi-squared test (n = 1556, p-value <0.001). The CHWs in the treatment group see sanitation marketing as their main job more than those in the no-treatment group.

Table 5 - Main Job

	No treatment	Treatment
0	210 (28%)	132 (16%)
1	546 (72%)	668 (84%)
Observations (Missings)	756 (0)	800 (0)

Budgetary Constraints

Furthermore, the variable Earnings (earn_san) was created to report payment of CHWs, reflecting their potential budgetary constraints. This variable reflects whether or not the marketer gets paid or not. 0 refers to not receiving any payment for the sanitation marketing, while 1 represents any

form of payment other than nothing. We do not distinguish between different wages. The variable was created as a combination of the question "How are you paid for sanitation marketing" with the possibility of specifying no payment in further questions. The selection of this variable to reflect budgetary constraints as well as detailed variable creation is available in the Appendix.

Table 6 - Earnings

	No treatment	Treatment
0	417 (47%)	132 (16%)
1	339 (53%)	668 (84%)
Observations (Missings)	756 (0)	800 (0)

To see how the answers of the variable were distributed, we employed descriptive statistics. The area that did not receive treatment had more unpaid CHW marketers than the area that received treatment. The CHWs operating in the area that received sanitation marketing treatment also had more CHWs that were paid, although overall more CHWs were paid than not paid. Conducting the Chi-squared test showed that the difference in the proportions of answers (n = 1556, p-value <0.001), thus the proportions of paid and unpaid CHW marketers, between the treatment and no treatment group were significant. Indeed, more CHW marketers are paid in the treatment area compared to those in the no treatment area.

Analysis - Inferential Statistics

Our analysis aims to examine whether the presence of sanitation marketing increases sanitation-related outcomes: hygiene behavior and the change of a sanitation system from the default one (meaning that the household installs, improves, or performs both modifications of a sanitation system). Before performing a regression analysis, a strong effect for both outcomes was already suggested in the descriptive statistics. The comparison of means for the first outcome variable – hygiene behavior – indicated that households that received sanitation marketing treatment showed more hygienic behavior as reported by the interviewer than those who did not receive marketing treatment. Similarly, performing Chi-squared test for the other outcome variable has already indicated that the households located in the area that did not receive sanitation marketing treatment had significantly higher rate of keeping their sanitation systems intact than those in the area that received sanitation marketing, meaning that they neither installed nor improved their sanitation system. Consequently, the promising descriptive observations of an effect led to building a regression models 1 and I to test hypothesis 1a and 1b respectively.

Different regression types were used depending on the character of the outcome variable. For regression analysis 1, an ordinary least squares regression (OLS) was used as the outcome hygiene behavior is a quantitative variable, while probit binary regression was used for regression analysis I as change of a sanitation system is a binary variable. The predictors in the models were introduced step-wise for both regressions, starting with area, then adding barazas, and then adding CHW characteristics.[5]

---

[5] We also conducted testing of interaction effects of each independent variable with area independent variable, which is available in the Appendix

Regression Analysis 1

Following the significant suggestions provided by the descriptive statistics, regression analysis was employed to test for hygiene behavior as an outcome variable. The first part of the regression analysis examines the impact of altogether 5 predictors on hygiene behavior, which is one of the two main sanitation outcomes measured in this study. In order to test for the effects of these predictors, OLS regression was used to build three main linear models as we assume that the relationship between the outcome variable hygiene behavior and the predictors is linear. The first model (1.1) aimed to examine if receiving sanitation marketing treatment increases hygiene behavior. The second model (1.2) aimed to examine if receiving sanitation marketing treatment as opposed to not receiving treatment and the number of health barazas influence hygiene behavior. Furthermore, the third model (1.3) analyzes the influence of the two predictors mentioned earlier and the CHW characteristics: time spent on sanitation marketing, seeing sanitation marketing as the main job compared to not seeing it as one´s main job, and whether the marketer gets paid for the job as opposed to volunteer work. [6]

All of the models use a variation of the basic linear regression equation for a quantitaive outcome variable: $Y = b_0 + b_1*area + b_2*barazas + b_3*CHW + e$. Y stands for the outcome variable, which is hygiene behavior in this case. b_0 depicts the constant of the regression, which is the Y-intercept of the regression model. b_1, b_2 and b_3 represent coefficients of independent variables. In this case, b_1 is the coefficient of the first predictor variable area, b_2 of the second predictor variable barazas etc. e stands for residual error term, representing the unmeasured variables. In our regression analysis, new predictors are introduced to the regression step-wise, meaning that the first regression model uses only the first predictor – area, second one uses both area and barazas, and the third model uses also CHW characteristics as 3 addictional predictors. For each regression model, the coefficients are depicted in Figure 1 with their level of significance.

As can be seen in Figure 1, Model 1 was calculated to to predict hygiene behavior (san_comb) based on whether or not the area received sanitation marketing treatment. It indicates that area is a significant predictor of hygiene behavior (n = 1397, p<.01). This means that those areas that received sanitation marketing did show in fact better hygiene behavior than those who did not receive it, which is in line with our descriptive results.

Moreover, the second regression model (1.2), also shown in the Figure 1, was calculated to predict hygiene behavior based on the number of barazas that held place and whether the area received sanitation marketing treatment as opposed to not receiving it. Controlling for the number of barazas does not change the significance level of area. Yet, the coefficient of barazas is significant and negative (n = 1397, p<.01). The number of barazas organized by the local CHW actually decreases hygiene behavior of the households. Yet, it is important to note that even though the effect of barazas was slightly negative, the coefficient lies within a wide confidence interval that reaches to almost neutral values. Therefore, barazas still significantly decrease hygiene behavior, but only with a slight decrease.

Thirdly, the third regression model (1.3) was calculated, building on the previous two models. The regression 1.3, depicted in Figure 1, shows that the area and barazas predictors remain significant also when controlling for CHW characteristics (n = 1397, p < .01). However, the CHW determinanats

---

[6] Additionally, an ordered logistic model was used analyzing the predictors used in regression (1.3), with hygiene behavior defined as a categorical ordinal variable with 3 levels. This ordinal logistic regression is reported in the Appendix as an expansion on the current OLS model 1.3.

themsleves do not show any significance. The coefficients of both barazas and area stayed very close to their previous values throughout the step-wise regression model building. Similarly, the constant did not change much throughout the three consecutive models, indicating that newly added variables do not interact with each other significantly.

	Dependent variable:		
	Hygiene Behavior		
	(1)	(2)	(3)
Area	0.109***	0.106***	0.129***
	(0.065, 0.153)	(0.062, 0.150)	(0.082, 0.176)
Number of barazas		-0.013**	-0.013**
		(-0.024, -0.002)	(-0.024, -0.002)
Time spent on marketing (%)			0.011
			(-0.004, 0.026)
Main job			-0.013
			(-0.081, 0.054)
Earnings			-0.047
			(-0.108, 0.014)
Constant	1.459***	1.506***	1.493***
	(1.428, 1.491)	(1.457, 1.556)	(1.397, 1.589)
Observations	1,397	1,397	1,397
R^2	0.017	0.021	0.027
Adjusted R^2	0.016	0.019	0.023
Residual Std. Error	0.419 (df = 1395)	0.418 (df = 1394)	0.417 (df = 1391)
F Statistic	23.780*** (df = 1; 1395)	14.797*** (df = 2; 1394)	7.661*** (df = 5; 1391)
Note:			*p**p***p<0.01

Figure 1 - Hygiene Behavior Regression Analysis

Regression Analysis II

The second important part of this research is seeing how sanitation marketing and CHW characteristics predict latrine installment or improvement as seen in the predicted outcome variable – change of sanitation system, which was reported by the locals themselves. There are altogether 5 independent variables that are used for 3 regression models analogous to those in regression analysis I. However, since the outcome variable is binary (0 indicating no change of a system and 1 indicating a positive change), probit regression models are used in this part of regression analysis, for which the z statistic was used for evaluation of significance of the coeffecents. All of the models use a variation of the probit regression, in which the results indicated in the Figure 2 are not intrepreted as coefficients indicating a direct effect, as was the case of OLS regression. Rather, Table 7 indicates the average of marginal

effects of the sample for each regression, that show how each variable affects the probability of the sanitation outcome occuring. [7]

Figure 2 desplays the first probit regression model (I.I) that was calculated to predict whether a household installs or improves their sanitation system as opposed to nothing based on whether it received sanitation marketing or not. We find that households' predicted change of their latrine facility was significantly predicted by area receiving sanitation marketing treatment (n = 1391, p < .01). The probability of a household changing a sanitation system increases when it is located in the area that received sanitation marketing.

Similarly, the regression model (I.II) was claculated to predict the change of a sanitation system based on area, while adding the number of barazas to the regresison model. We find that households´ predicted change of sanitation system was significantly predicted by both area receiving treatment as opposed to not receiving treatment and the newly-added number of barazas, which had a negative effect (n = 1391, p < .01). While the probability of change of a sanitation system increases with household being located in a treatment area as opposed to no treatment area (just as in regression I.I), it decreases with the number of barazas that took place.

Moreover, the regression model (I.III) was claculated by adding CHW characteristics to the previous model. While area and barazas remain significant with almost unchanged average of sample marginal effects, none of the CHW characteristics shows a significant effect on the change of a saniation system. Moreover, the constant of this regression was not significant either, indicating that there was too much noise within the constant.

---

[7] Stemming from description analysis, we conducted interaction testing in the Appendix as well as an OLS version of the regression model I (change of a sanitation system outcome variable)

		Dependent variable:	
		Change of Sanitation System	
	(I)	(II)	(III)
Area	0.214***	0.204***	0.209***
	(0.081, 0.347)	(0.071, 0.337)	(0.065, 0.352)
Number of barazas		-0.048***	-0.050***
		(-0.080, -0.015)	(-0.083, -0.017)
Time spent on marketing (%)			0.022
			(-0.023, 0.067)
Main job			0.079
			(-0.126, 0.284)
Earnings			-0.009
			(-0.193, 0.175)
Constant	0.091*	0.262***	0.121
	(-0.003, 0.186)	(0.110, 0.413)	(-0.171, 0.414)
Observations	1,391	1,391	1,391
Log Likelihood	-941.579	-937.547	-936.932
Akaike Inf. Crit.	1,887.158	1,881.095	1,885.863
Note:			*p**p***p<0.01

Figure 2 - Change of Sanitation System Regression Analysis

	Regression I	Regression II	Regression III
(Intercept)	0.036*	0.101***	0.047
Area	0.083***	0.079***	0.08**
Number of barazas		-0.018***	-0.019**
Time spent on sanitation marketing (%)			0.008
Main Job			0.03
Earnings			-0.004

Table 7 - Averages of sample marginal effects for each regression

Discussion

The discussion section presents answers to each research question by evaluating results of our study. We also consult the results with existing academic literature to provide us with an outlook on our results. Each question is analyzed from the point of both sanitation outcomes, finding parallels in them and distinguishing differences between them.

Research Question 1

The results of our study clearly show that sanitation marketing has a positive impact on both hygiene behavior of the households and the change of their sanitation system as opposed to keeping an old one. When consulting the results with academic literature, the significant positive impact of sanitation treatment is not surprising, although some studies found mixed results in the past (Evans et al., 2014; Grier & Bryant, 2005; Jenkins & Scott, 2007). The finding that sanitation marketing improved both outcomes also suggests that the change of sanitation system and behavioral change might be closely intertwined, as having a sufficient system is an important prerequisite for good hygiene behavior (Jenkins & Scott, 2007). Additionally, since reporting the change of a sanitation system could have been quite subjective, it is an important finding since both variables had similar results in the end, indicating that the locals probably indicated accurate numbers. Therefore, both hypotheses 1a and 1b are supported.

As Grier and Bryant have noted, social marketing is a highly effective strategy in improving sanitation and societal goals (Grier & Bryant, 2005), in our case it was sanitation improvement seen at a two-fold level of improving a sanitation system and corresponding hygiene behavior. This is mostly due to social marketing being led by the community process, identifying its specific goals and thus allowing the marketers to use the most fitting strategy to sell the product (Grier & Bryant, 2005). Sanitation marketers, who mostly operate in their own village are highly familiar with the locals and can direct their strategy to advertise the product as to fit the needs and wants of the local household. The results illustrate that one-on-one marketing at the place of living by members of a local community – marketers, can be efficient as CHW marketers can readily adopt their strategy to each household´s needs. Our results are also in line with literature findings that emphasize the relation of CHWs and locals as crucial when achieving successful health results (Jenkins & Curtis, 2005).

The success of sanitation marketing is significant despite many factors that could have potentially diminished its effect. Successful sanitation marketing needs to overcome several difficulties that can prevent locals from adopting a better sanitation system and improving hygiene behavior. As studies point out, achieving better health is oftentimes not the actual motivator in improving sanitation, but it is rather lowering costs associated with better sanitation (Jenkins & Curtis, 2005; Rosenboom, Jacks, Phyrum, Roberts, & Baker, 2011). Direct door-to-door sanitation marketing proves to be effective in minimizing the cost of engagement for the locals (Grier & Bryant, 2005; Rosenboom et al., 2011). Villagers did not have to leave their households to inform themselves about the program as they were approached directly. Minimizing costs in general, including those associated with time and energy when learning about sanitation, is a very important factor of success of sanitation marketing (Jenkins & Curtis, 2005). Therefore, our results suggest that the villagers who could have been prevented from installing a new sanitation system only by not being informed or approached before, adopted better sanitation systems in the face of direct, cost-minimizing sanitation marketing.

Research Question 2

Not only did this study examine the effect of sanitation marketing alone, it was also analyzing what determines the success of sanitation marketing from the side of the CHW marketer and the impact of health barazas through which these characteristics can be manifested. We aimed to study CHW´s status, motivation and budgetary constraints as the factors determining CHW´s performance. As CHW´s performance was measured by the same sanitation outcomes on the household level as were used for studying the overall effect of sanitation marketing, the regression models used for answering research question 1 were further refined to answer the research question 2 and address the second part of hypotheses posed in the beginning of the paper.

The impact of barazas was analyzed as it was expected to act both as a variable indicating intrinsic motivation of a CHW and a potential mediating variable, increasing people´s exposure to the social marketing ideas as hypothesized. However, the results for barazas prediction were surprising. Namely, the number of health barazas organized by a CHW decrease the actual sanitation outcome in both forms – hygiene behavior and change of a sanitation system. This indicates a possible dissociation of what actions are taken by the local CHW and whether or not the villagers in the area change their sanitation habits. Therefore, the hypothesis 2a is rejected.

Comparing the negative impact of barazas to direct sanitation marketing is especially interesting. While in-person approach that is used in sanitation marketing encourages households to both change a sanitation system for the better and improve their respective hygiene behavior, organization of health barazas does not have the same effect. As studies suggest, one of the main reasons a household decides to install or improve its sanitation system is lowering the costs associated with the change as much as possible (Jenkins & Scott, 2007). While this is done successfully in sanitation marketing as our results show, attending health barazas costs far more energy and other resources than learning about good sanitation practices and facilities at the villager´s household. Another factor that also causes low attendance of community meetings is that they are usually attended only by villagers of a higher status, which means that the less wealthy were not exposed to the potential effect of baraza meetings (Ghazala & Rao, 2004). Simply said, the barazas could have been organized by a motivated CHW, yet they probably were not attended by a large portion of villagers as it costs them energy and time resources or because they were of a lower social status. On the other hand, a much larger number of villagers were reached through sanitation marketing performed on a door-to-door basis. This result also ties into the study by Rosenboom and colleagues who found that latrine adoption must be associated with selling behavior, rather than information provision (Rosenboom et al., 2011). Rosenbloom´s results suggest that even if enough people attended the health barazas, what would make a difference is still the persuasive activity associated with selling behavior (Rosenboom et al., 2011).

Furthermore, we also analyzed how well different CHW characteristics predict the villagers´ sanitation outcomes. However, contrary to the hypotheses 2b, 2c and 2d, the CHW predictors were not significant for predicting either hygiene behavior or change of a sanitation system as an outcome.

Overall, this can be explained due to a wide range of other factors that affect sanitation outcomes. As Jenkins and Scott note, there are multiple factors affecting whether or not the household decides to adopt a new sanitation system in the end (Jenkins & Scott, 2007). Among many, the priority of a sanitation system change must be higher than other competing goals (Jenkins & Scott, 2007). Moreover, there cannot be any permanent or temporary constraints that would prevent the household from installing a new system and changing its hygiene behavior (Jenkins & Scott, 2007). These factors cannot be influenced by CHW marketers sufficiently as it is not in their capability to

address for example lacking financial resources of a household that would want to improve its sanitation habits. What CHW marketers could potentially influence is the prioritization of improving a household´s sanitation system, yet this can only work to an extent. Regardless of how motivated, prestigious or skilled a CHW is to motivate a household to place sanitation improvement as a higher priority, CHWs cannot address many of the constraints a household faces to adopt a better system. It is probable that the lower prioritization is not the most important factor blocking the actual sanitation improvement, as this could be addressed by a highly motivated CHW. It appears that the constraints, most likely financial resources, are the main factor not allowing locals to change their sanitation system and thus behavior remains unchanged as well. More research is needed to analyze whether villagers view CHW´s characteristics as an important factor when making a decision to improve their sanitation system.

However, when looking at the predictors separately, it is also possible that it is not that overall CHW characteristics do not influence the sanitation improvements, but that these particular characteristics are not good predictors of CHW characteristics. Yet, this would be rather surprising in light of literature, especially for receiving or not receiving payment. Multiple studies have indicated that CHWs cite the lack of financial resources as one of the main problems not allowing them to perform well in their job (Aseyo et al., 2018; Olayo et al., 2014). This was especially the case of community health volunteers in the study conducted by Aseyo, in which community health volunteers indicated that receiving a monetary compensation would increase their motivation for the job and improve the health results of their communities (Aseyo et al., 2018). Despite this claim, our study finds that not only receiving payment, as opposed to not receiving a payment, does not improve sanitation results of communities, but not even time spent on sanitation marketing was found to improve sanitation outcomes.

This is a rather surprising result that needs to be also explained with other constraining factors that do not permit households to improve their sanitation systems. As studies found, motivation is one of the best predictors of the performance of a CHW (Ashraf et al., 2020), respectively the lack of it prevents them from achieving good results (Aseyo et al., 2018). Yet, these studies need to be viewed as context-specific. Despite Ashraf reporting various health improvements and more engagement with one´s work for highly pro-socially motivated CHWs, the study was conducted in Zambia and the main constraints faced by the local population were different than in case of our study (Ashraf et al., 2020). Therefore, our study finds that motivation cannot compensate for the lacking resources local households and communities have. Moreover, what is even more surprising is that the percentage of time that the CHW spent on sanitation marketing is not associated with better sanitation outcomes alone. This is surprising as Besley and Ghatak suggest that this work should be aligned with both high motivation and higher status (Besley & Ghatak, 2005).

Lastly, unlike other studies, we do not find a relation of a CHW´s status reported through seeing sanitation marketing as a main job either. Several studies found social prestige to be associated with better CHW performance (Alam et al., 2012; Y. Kawakatsu et al., 2012; Tesfay, Yemane, & Gebreselassie, 2014). However, similarly to another study by Kawakatsu, who reported the effect of social capital as not significant for a CHW performance (Kawakatsu et al., 2015), we find comparable results. Social status associated with seeing sanitation marketing as one´s main job does not increase sanitation outcomes in the community. However, similarly to other CHW characteristics, the weak predictive power can be due to other constraints that prevent locals in the particularly studied counties from adopting a sanitation system.

Limitations

This study has several limitations that need to be kept in mind when generalizing the results. Among others, this study was just an observational study. There can be confounding variables that were not accounted for, since the areas can have certain particularities that could have influenced the results. One of them is the state of sanitation prior to the marketing treatment that could have shifted the results. It is possible that some villages had already better sanitation systems or that they were employing better sanitation practices. In addition, there was a problem with the distribution of observations within no treatment constituency Bongoma, which were not equally distributed between the two sub-areas in the constituency. There were more villages located in Sirisia than in Bumula, thus the sample of no treatment villages was not entirely hegemonic. Moreover, as highlighted in the discussion, one of the greatest limitations of the study is not accounting for other factors influencing sanitation outcomes from the villagers´ perspective that could have diverted the results regrading CHW characteristics. This was also demonstrated by a relatively low predictability of our regression models. Future studies are encouraged to research CHW performance directly in form of their outputs, not just as a health outcome. When health outcomes are studied, it is important that future studies address and allow locals to report why the adoption of a health improvement was either successful or unsuccessful.

Furthermore, one of the greatest limitations that could have influenced the results are spillover effects and similar sanitation marketing programs happening in the no treatment area, thus not allowing for a perfect comparison. As no-treatment group of CHWs also reported sanitation marketing activities, it is obvious that other similar programs were taking place parallel to the program that was studied. This practically means that we could have studied only the particular program provided by FINISH INK´s partner organizations. Moreover, it is possible that some marketers were migrating from one area to another, living in a very close proximity to areas from a different group and thus either receiving or not receiving the training, as opposed to what was intended. Future study design should rather be directed at measurement of sanitation outcomes at several time points, using the same sample of villages and households rather than comparison.

Furthermore, many variables used in the study were rather subjective and could have been influenced by the recall effect, in which participants reported better outcomes or more barazas happening than was really the case. The study tried to compensate for problematic measurement of improvement or installation of a new sanitation facility by including sanitation behavior that was reported by the interviewer. Yet, also the interviewer could have been influenced by many factors and this variable is therefore not completely strategy-proof either. It is also important to point out that hygiene behavior could have been also measured separately as sanitation systems usage and its maintenance. These two practices seem related, yet future studies should also compare these different components of hygiene behavior when evaluating the effect of sanitation marketing.

Conclusion

Sanitation marketing remains a widely used practice to improve health in many countries. Our study examined the effect of sanitation marketing on households located in Western Kenya, where sanitation marketing concentrated on improving sanitation systems and related hygiene behavior. Moreover, we studied the characteristics of local CHW marketers that deliver the program to their communities. By analyzing time they spend on sanitation marketing, whether they see sanitation marketing as their main job, and whether they get paid for it, we analyzed the relation of their motivation, status and budgetary constraints with the sanitation outcomes of the villagers they were

responsible for. The study also analyzed the relation of health barazas frequency with the sanitation outcomes of West Kenyans. As we found, sanitation marketing is an important and effective way to address the problems of insufficient sanitation.

Many programs already use this practice, and our study provides supportive evidence for it. However, we did not find any evidence for the studied CHW characteristics having any impact on hygiene behavior nor on improving sanitation systems of the locals. What appears to be the successful method in using sanitation marketing is lowering costs associated with adoption of new sanitation habits. Yet, as we found, sanitation marketing can lower the costs only to a limited extent. CHWs are unable to motivate change after the marginal costs are overcome by basic sanitation marketing, even though they can be highly motivated, hold a high status and be receive an adequate financial compensation. We suggest that there are certain constraint households face that cannot be overcome by sanitation marketing. Opposite to the finding that sanitation marketing lowers energy and time costs associated with learning about new sanitation methods, health barazas might actually increase the costs and thus be negatively associated with sanitation outcomes of the locals. This research encourages further studies of the effect of health barazas as well as CHW characteristics on sanitation outcomes, as well as analyzing the other side – villagers. It is crucial to study what motivates and enables people to adopt better sanitation systems and behaviors and how they perceive the CHW characteristics in that matter. Social marketing is a growing in its popularity, and it is doing so rightfully.

Bibliography

Alam, K., Tasneem, S., & Oliveras, E. (2012). Performance of female volunteer community health workers in Dhaka urban slums. *Soc Sci Med, 75*(3), 511-515. doi:10.1016/j.socscimed.2012.03.039

Aseyo, R. E., Mumma, J., Scott, K., Nelima, D., Davis, E., Baker, K. K., . . . Dreibelbis, R. (2018). Realities and experiences of community health volunteers as agents for behaviour change: evidence from an informal urban settlement in Kisumu, Kenya. *Human Resources for Health, 16*(1), 53. doi:10.1186/s12960-018-0318-4

Ashraf, N., Bandiera, O., Davenport, E., & Lee, S. S. (2020). Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. *American Economic Review, 110*(5), 1355-1394. doi:10.1257/aer.20180326

Ballard, M., Madore, A., Johnson, A., Keita, Y., Haag, E., Palazuelos, D., & Rosenberg, J. (2018). Community Health Workers.

Besley, T., & Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *The American Economic Review, 95*(3), 616-636.

Edward, A., Jung, Y., Chhorvann, C., Ghee, A. E., & Chege, J. (2019). Association of mother's handwashing practices and pediatric diarrhea: evidence from a multi-country study on community oriented interventions. *Journal of preventive medicine and hygiene, 60*(2), E93-E102. doi:10.15167/2421-4248/jpmh2019.60.2.1088

Ellis, A., McClintic, E. E., Awino, E. O., Caruso, B. A., Arriola, K. R. J., Ventura, S. G., . . . Freeman, M. C. (2020). Practices and Perspectives on Latrine Use, Child Feces Disposal, and Clean Play Environments in Western Kenya. *Am J Trop Med Hyg, 102*(5), 1094-1103. doi:10.4269/ajtmh.19-0389

Evans, W. D., Pattanayak, S. K., Young, S., Buszin, J., Rai, S., & Bihm, J. W. (2014). Social marketing of water and sanitation products: A systematic review of peer-reviewed literature. *Social Science & Medicine, 110*, 18-25. doi:https://doi.org/10.1016/j.socscimed.2014.03.011

Ghazala, M., & Rao, V. (2004). Community-Based and -Driven Development: A Critical Review. *The World Bank Research Observer, 19*(1), 1-39.

Grier, S., & Bryant, C. A. (2005). SOCIAL MARKETING IN PUBLIC HEALTH. *Annual Review of Public Health, 26*(1), 319-339. doi:10.1146/annurev.publhealth.26.021304.144610

Jenkins, M. W., & Curtis, V. (2005). Achieving the 'good life': Why some people want latrines in rural Benin. *Social Science & Medicine, 61*(11), 2446-2459. doi:https://doi.org/10.1016/j.socscimed.2005.04.036

Jenkins, M. W., & Scott, B. (2007). Behavioral indicators of household decision-making and demand for sanitation and potential gains from social marketing in Ghana. *Soc Sci Med, 64*(12), 2427-2442. doi:10.1016/j.socscimed.2007.03.010

Kawakatsu, Y., Sugishita, T., Kioko, J., Ishimura, A., & Honda, S. (2012). Factors influencing the performance of community health workers in Kisumu West, Kenya. *Prim Health Care Res Dev, 13*(4), 294-300. doi:10.1017/s1463423612000138

Kawakatsu, Y., Sugishita, T., Tsutsui, J., Oruenjo, K., Wakhule, S., Kibosia, K., . . . Honda, S. (2015). Individual and contextual factors associated with community health workers' performance in Nyanza Province, Kenya: a multilevel analysis. *BMC Health Services Research, 15*(1), 442. doi:10.1186/s12913-015-1117-4

Kawakatsu, Y., Tanaka, J., Ogawa, K., Ogendo, K., & Honda, S. (2017). Community unit performance: factors associated with childhood diarrhea and appropriate treatment in Nyanza Province, Kenya. *BMC public health, 17*(1), 202. doi:10.1186/s12889-017-4107-0

Naanyu, V., Sidle, J. E., Frankel, R. M., Ayuku, D., Nyandiko, W. M., & Inui, T. S. (2011). Rooting inquiry in tradition: the health baraza as a tool for social research in Kenya. *Qual Health Res, 21*(1), 14-26. doi:10.1177/1049732310367498

Njomo, D. W., Masaku, J., Mwende, F., Odhiambo, G., Musuva, R., Matey, E., . . . Kihara, J. H. (2017). Local stakeholders' perceptions of community sensitization for school-based deworming

programme in Kenya. *Tropical diseases, travel medicine and vaccines, 3*, 15.
doi:10.1186/s40794-017-0058-9

Ojakaa, D., Okoth, E., Wangila, S., Ndirangu, M., Mwangi, N., & Ilako, F. (2011). Making aid effective
at the community level: the AMREF experience. *Development in Practice, 21*(7), 1013-1022.

Olayo, R., Wafula, C., Aseyo, E., Loum, C., & Kaseje, D. (2014). A quasi-experimental assessment of
the effectiveness of the Community Health Strategy on health outcomes in Kenya. *BMC
Health Services Research, 14 Suppl 1*, S3. doi:10.1186/1472-6963-14-S1-S3

Rosenboom, J. W., Jacks, C., Phyrum, K. O. V., Roberts, M., & Baker, T. (2011). Sanitation marketing in
Cambodia. *Waterlines, 30*(1), 21-40.

Tesfay, F., Yemane, D., & Gebreselassie, A. (2014). Assessment of non-financial incentives for
volunteer community health workers – the case of Wukro district, Tigray, Ethiopia. *Human
Resources for Health, 12*, 54. doi:10.1186/1478-4491-12-54

Wangalwa, G., Cudjoe, B., Wamalwa, D., Machira, Y., Ofware, P., Ndirangu, M., & Ilako, F. (2012).
Effectiveness of Kenya's Community Health Strategy in delivering community-based maternal
and newborn health care in Busia County, Kenya: non-randomized pre-test post test study.
The Pan African medical journal, 13 Suppl 1, 12.

Appendix 1

House_lat_comb: Hygiene behavior dependent variable creation

Questions concerning this variable were:

house_lat_use_signs	**TO INTERVIEWER: Are there signs of usage (such as a footpath from the house to the latrine)?**
House_lat_clean	**TO INTERVIEWER: Does the latrine seem clean and well-maintained?**
House_lat_use_wash	TO INTERVIEWER: Are there hand washing facilities?
House_lat_use_soap	TO INTERVIEWER: Is there soap?

Although it was initially intended to combine relevant answers from all of the relevant questions above, in the end due to missing data and incompatible coding of answers, only the two most relevant outcomes were chosen. Namely, house_lat_use_signs and house_lat_clean were preferred over washing facilities and the presence of soap

	House_lat_clean	House_lat_use_signs	House_lat_comb
Min.	0	0	0
1st Qu.	1	2	1.5
Median	1	2	1.5
Mean	1.236	1.796	1.516
3rd Qu.	2	2	2
Max.	2	2	2
NA´s	159	159	159

Appendix 1

Appendix 1 depicts the creation of the outcome variable house_lat_comb that acts as one of the main sanitation outcome variables. The overall variable represents hygiene behavior. However, we note that there might be an advantage to see how the variables comprised in the overall hygiene behavior compare and how the main variable was created. House_lat_clean represents cleanness and good maintenance of a sanitation system, while house_lat_use_signs represents whether the system seems frequently used which can be visible for example by a foot pathway leading to the sanitation system. We assume that these two variables are strongly correlated, thus they can be combined into one overall variable house_lat_comb, which is indeed the case as tested for by Pearson´s correlation test ($t = 9.6716$, $df = 1395$, p-value $< 2.2e-16$, $cor = 0.250679$). It is assumed that good hygiene behavior comprises of both components – cleaning and maintain the facility as well as the facility being used often. Nevertheless, for more detailed information, further appendix involves regression analysis of the two outcomes separately as well.

House_lat_new dependent variable creation

Locals answered the question "Have you newly built or improved this latrine since 2015?". The categorical variable has three ordinal levels representing answers of:

0: no
1: yes, improved
2: yes, build new
3: yes, it was newly constructed and then further improved
9999: I do not know/ I do not remember

The 9999 answers were recoded as missing observations as they do not provide us with any analyze-able content. Levels 1,2 and 3 were jointly coded as 1 as they all include changing the current sanitation system of a household, thus standing for improvement of some sort – be it installment of a new latrine, improving an old one or both. The 0 level stayed unchanged.

The variable house_lat_new was originally coded in different levels, yet it was coded as a dummy binary variable for the simplicity of regression analysis interpretation. Here, in the figure Appendix 2, we presnet the descriptive statistics of the original variable for greater clarity. Originally, there were 81 and 78 missing answers in the respective groups and 6 households reporting that they did not remember or did not know. The answers coded as 9999 depicting not knowing the answer were also coded as missings afterwards.

In order to test if there are significant differences between households in regions that received sanitation marketing treatment and those who did not, we performed the non-parametric Pearson´s Chi-squared test. The test showed a significant difference between latrine installment in households that received sanitation marketing (X-squared = 15.881, df = 3, p = 0.001199). The figure Appendix 3 shows Chi-square residuals, which are the differences in specific categories that differ between the two groups.

Appendix 3

Characteristic	no_treatment, N = 756[1]	treatment, N = 800[1]
house_lat_new		
0	312 / 675 (46%)	273 / 722 (38%)
1	30 / 675 (4.4%)	58 / 722 (8.0%)
2	326 / 675 (48%)	376 / 722 (52%)
3	5 / 675 (0.7%)	11 / 722 (1.5%)
9999	2 / 675 (0.3%)	4 / 722 (0.6%)
(Missing)	81	78

[1] Statistics presented: n / N (%)

Appendix 2- Chi-square residuals

	no_treatment	treatment
0	1.722	-1.667
1	-1.927	1.866
2	-0.74	0.717
3	-0.985	0.954

Job_san independent variable creation

The variable job_san was recoded into the levels 0 and 1 corresponding to the answers 2 and 3 respectively. Appendix 4 depicts the original variable distribution in the old coding. Appendix 5 presents the Chi-square residuals after performing the Chi-squared test.

Characteristic	no_treatment, N = 756[1]	treatment, N = 800[1]
job_san		
2	546 / 756 (72%)	668 / 800 (84%)
3	210 / 756 (28%)	132 / 800 (16%)

[1] Statistics presented: n / N (%)

Appendix 4

	no_treatment	treatment
0	3.401	-3.306

1	-1.805	1.755

Appendix 5

Earn_san: Earnings independent variable creation

Earnings variable (earn_san) was created by combining several variables asking for how and how much CHW marketers are paid, using the following questions:

earn_san   How are you paid for sanitation marketing?

earn_san_oth   Other, specify:

earn_san_amount_hh   How much are you paid per household?

earn_san_amount_mnth   What is your fixed rate per month?

earn_job_mnth   How much are you paid for your main job per month?

While the other questions asked to report a specific number as a wage as a follow-up to the first question (with the possibility to report 0, if they are not paid), the possible answers for the first question, "How are you paid for sanitation marketing?" were:

1	Per household that you approach
2	Per household that you refer to a financial institution
3	Per household that installs a sanitation system
4	Fixed rate
5	As part of my main job
6	Other

Based on which option the CHW chose, a corresponding further question was asking him to report more details on his/her payment. If the CHW chose option 6 – other, he/she could report further details in the variable earn_san_oth. Most of the answers reported here, were marketers reporting not receiving payment in various wordings, such as "volunteer" or "not paid". All of these were recoded as 0. Few observations reported small allowance or being paid as a part of their training, these were coded as 1.

Moreover, other follow-up questions to the first question were examined. For all the other variables, the observations reporting 0 did not change their coding. However, all the numerical values higher than 0, thus indicating that the CHW is paid, were coded as 1. In the end, we checked that all of those who reported not receiving payment as a part of the earn_san_oth and those who reported 0 in other questions, were truly assigned 0 in the new variable.

Hygiene Behavior O-logit model

In order to gain greater detail when analyzing Hygiene behavior, we also performed an O-logit regression, since Hygiene behavior variable could be also seen as a categorical variable with different ordinal levels. Appendix 6, Model 4 depicts the O-logit model as a comparison to the three main OLS models used in the main analysis.

	Dependent variable:			
	Hygiene Behavior O-logit Regression *OLS*			house_lat_comb *ordered logistic*
	(1)	(2)	(3)	(4)
Area	0.109***	0.106***	0.129***	0.631***
	(0.065, 0.153)	(0.062, 0.150)	(0.082, 0.176)	(0.418, 0.843)
Number of barazas		-0.013**	-0.013**	-0.061**
		(-0.024, -0.002)	(-0.024, -0.002)	(-0.109, -0.013)
Time spent on marketing (%)			0.011	0.055*
			(-0.004, 0.026)	(-0.010, 0.121)
Main job			-0.013	-0.063
			(-0.081, 0.054)	(-0.362, 0.236)
Earnings			-0.047	-0.217
			(-0.108, 0.014)	(-0.485, 0.050)
Constant	1.459***	1.506***	1.493***	
	(1.428, 1.491)	(1.457, 1.556)	(1.397, 1.589)	
Observations	1,397	1,397	1,397	1,397
R^2	0.017	0.021	0.027	
Adjusted R^2	0.016	0.019	0.023	
Residual Std. Error	0.419 (df = 1395)	0.418 (df = 1394)	0.417 (df = 1391)	
F Statistic	23.780*** (df = 1; 1395)	14.797*** (df = 2; 1394)	7.661*** (df = 5; 1391)	
Note:				*p**p***p<0.01

Appendix 6

Change of Systems OLS models

Similarly to employing a different type of analysis for greater detail and comparison of regression models in the case of Hygiene behavior, we also used OLS regression to model for the outcome variable Change of Systems. Appendix 7, Model 4 depicts the OLS regression done by the same variables as were used in the main analysis for probit regression.

	Dependent variable:			
	Installment or Improvement of the Sanitation System (OLS)			
	probit			OLS
	(I)	(II)	(III)	(4)
Area	0.214***	0.204***	0.209***	0.081***
	(0.081, 0.347)	(0.071, 0.337)	(0.065, 0.352)	(0.026, 0.137)
Number of barazas		-0.048***	-0.050***	-0.019***
		(-0.080, -0.015)	(-0.083, -0.017)	(-0.032, -0.006)
Time spent on marketing (%)			0.022	0.009
			(-0.023, 0.067)	(-0.009, 0.026)
Main job			0.079	0.031
			(-0.126, 0.284)	(-0.049, 0.111)
Earnings			-0.009	-0.004
			(-0.193, 0.175)	(-0.076, 0.068)
Constant	0.091*	0.262***	0.121	0.548***
	(-0.003, 0.186)	(0.110, 0.413)	(-0.171, 0.414)	(0.435, 0.661)
Observations	1,391	1,391	1,391	1,391
R^2				0.014
Adjusted R^2				0.010
Log Likelihood	-941.579	-937.547	-936.932	
Akaike Inf. Crit.	1,887.158	1,881.095	1,885.863	
Residual Std. Error				0.491 (df = 1385)
F Statistic				3.870*** (df = 5; 1385)
Note:				*p**p***p<0.01

Appendix 7

House_lat_use_signs model

Appendix 8 provides the results of a component of Hygiene behavior – usage of sanitation systems, when modelled with simple OLS regression using the same predictors as were used in the main analysis, to provide a greater detail of our analysis.

	Dependent variable:		
	Sanitation Usage Signs		
	(1)	(2)	(3)
Area	0.047**	0.045*	0.079***
	(0.023)	(0.023)	(0.025)
Number of barazas		-0.009	-0.008
		(0.006)	(0.006)
Time spent on marketing (%)			0.012
			(0.008)
Main job			-0.024
			(0.036)
Earnings			-0.078**
			(0.032)
Constant	1.772***	1.803***	1.805***
	(0.017)	(0.027)	(0.051)
Note:			*p**p***p<0.01

Appendix 8

House_lat_clean model

Similarly, Appendix 9 provides the results of another component of Hygiene behavior – cleanness and good maintenance of a sanitation system, when modelled with simple OLS regression using the same predictors as were used in the main analysis, to provide a greater detail of our analysis.

Maintatance and Cleanness Regression Results			
	Dependent variable:		
	Maintanance and Cleanness of a Sanitation System		
	(1)	(2)	(3)
Area	0.172*** (0.108, 0.236)	0.168*** (0.104, 0.232)	0.178*** (0.109, 0.248)
Number of barazas		-0.018** (-0.034, -0.002)	-0.018** (-0.034, -0.002)
Time spent on marketing (%)			0.009 (-0.012, 0.031)
Main job			-0.003 (-0.103, 0.096)
Earnings			-0.016 (-0.106, 0.073)
Constant	1.147*** (1.100, 1.193)	1.210*** (1.137, 1.283)	1.181*** (1.040, 1.322)
Observations	1,397	1,397	1,397
R^2	0.019	0.023	0.023
Adjusted R^2	0.019	0.021	0.020
Residual Std. Error	0.614 (df = 1395)	0.613 (df = 1394)	0.613 (df = 1391)
F Statistic	27.378*** (df = 1; 1395)	16.123*** (df = 2; 1394)	6.694*** (df = 5; 1391)
Note:			*p**p***p<0.01

Appendix 9

Appendix 2

Models with interactions – Hygiene behavior - house_lat_comb

Appendix 10 present each variable separately when interacted with the area variable in different models with Hygiene behavior as the outcome. After testing each variable separately, we built a joint interaction model depicted in Appendix 11, Model 4. In this model, we tested for area interactions with: barazas, main job, time spent and budgetary constrains in one model.

When testing for interaction, the model looks similar to the basic OLS model Y= b0 + b1*area + b2*barazas + b3*CHW characteristics +b4*area*barazas + b5*area*CHW characteristics + e. The difference from a basic regression model is that the effect of barazas and CHW chasracteristics is different for different areas. The unique effect of barazas is represented by everything that is multiplied by the variable in the model, therefore the effect is b2 + b4* barazas. The coeffeicient b2 is the effect of barazas on hygiene behavior only when area is the control group, while the coefficient of interacted term b4 captures the difference between the treated and control areas for the variable. The same interpretation goes for other interacted variables with their coefficients.

Appendix 11, Model 4 shows that Main job and Earnings were significant when interacted with area variable. The interacted term for Main job was positive, however it was negative for the interacted Earnings. Barazas is not significant anymore, although area remained a significant predictor.

				Dependent variable:				
				Hygiene Behavior				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Area	0.109***	0.106***	0.117***	0.129***	0.138***	0.164***	0.075	0.179***
	(0.065, 0.153)	(0.062, 0.150)	(0.072, 0.162)	(0.082, 0.176)	(0.051, 0.225)	(0.043, 0.286)	(-0.023, 0.173)	(0.086, 0.273)
Number of barazas		-0.013**	-0.013**	-0.013**	-0.010	-0.013**	-0.013**	-0.013**
		(-0.024, -0.002)	(-0.024, -0.002)	(-0.024, -0.002)	(-0.026, 0.006)	(-0.024, -0.002)	(-0.024, -0.002)	(-0.024, -0.002)
Time spent on marketing (%)			0.011	0.011	0.011	0.017	0.009	0.012
			(-0.004, 0.025)	(-0.004, 0.026)	(-0.004, 0.025)	(-0.004, 0.037)	(-0.005, 0.024)	(-0.003, 0.027)
Main job			-0.042	-0.013	-0.042	-0.037	-0.065*	0.00001
			(-0.099, 0.015)	(-0.081, 0.054)	(-0.099, 0.015)	(-0.095, 0.021)	(-0.141, 0.010)	(-0.071, 0.071)
Earnings				-0.047				
				(-0.108, 0.014)				
Area:Barazas (interaction)					-0.006			
					(-0.028, 0.016)			
Area: Time (interaction)						-0.012		
						(-0.040, 0.016)		
Area: Main job (interaction)							0.052	
							(-0.057, 0.162)	
Area: Earnings (interaction)								-0.028
								(-0.096, 0.039)
								(-0.197, 0.003)
Constant	1.459***	1.506***	1.490***	1.493***	1.479***	1.461***	1.512***	1.467***
	(1.428, 1.491)	(1.457, 1.556)	(1.394, 1.586)	(1.397, 1.589)	(1.374, 1.583)	(1.343, 1.579)	(1.406, 1.619)	(1.363, 1.572)
Observations	1,397	1,397	1,397	1,397	1,397	1,397	1,397	1,397
R^2	0.017	0.021	0.025	0.027	0.025	0.026	0.026	0.028
Adjusted R^2	0.016	0.019	0.022	0.023	0.022	0.022	0.022	0.024
Residual Std. Error	0.419 (df = 1395)	0.418 (df = 1394)	0.417 (df = 1392)	0.417 (df = 1391)	0.417 (df = 1391)	0.417 (df = 1391)	0.417 (df = 1391)	0.417 (df = 1390)

Appendix 10

	Dependent variable:			
	Hygiene Behavior			
	(1)	(2)	(3)	(4)
Area	0.109***	0.106***	0.129***	0.166*
	(0.065, 0.153)	(0.062, 0.150)	(0.082, 0.176)	(-0.022, 0.354)
Number of barazas		-0.013**	-0.013**	-0.010
		(-0.024, -0.002)	(-0.024, -0.002)	(-0.026, 0.006)
Time spent on marketing (%)			0.011	0.013
			(-0.004, 0.026)	(-0.009, 0.035)
Main job			-0.013	-0.056
			(-0.081, 0.054)	(-0.142, 0.030)
Earnings			-0.047	-0.007
			(-0.108, 0.014)	(-0.076, 0.062)
Area:Barazas (interaction)				-0.006
				(-0.027, 0.016)
Area: Time (interaction)				-0.007
				(-0.036, 0.023)
Area: Main job (interaction)				0.247***
				(0.081, 0.412)
Area: Earnings (interaction)				-0.246***
				(-0.403, -0.088)
Constant	1.459***	1.506***	1.493***	1.481***
	(1.428, 1.491)	(1.457, 1.556)	(1.397, 1.589)	(1.338, 1.623)
Observations	1,397	1,397	1,397	1,397
R^2	0.017	0.021	0.027	0.035
Adjusted R^2	0.016	0.019	0.023	0.029
Residual Std. Error	0.419 (df = 1395)	0.418 (df = 1394)	0.417 (df = 1391)	0.416 (df = 1387)
F Statistic	23.780*** (df = 1; 1395)	14.797*** (df = 2; 1394)	7.661*** (df = 5; 1391)	5.563*** (df = 9; 1387)

Note: *p**p***p<0.01

Appendix 11

Model with interactions – Change of sanitation system - house_lat_new

Appendix 12 present each variable separately when interacted with the area variable in different models with the Change of a sanitation system as the outcome. After testing each variable

separately, we built a joint interaction model depicted in Appendix 13, Model IV. In this model, we tested for area interactions with: barazas, main job, time spent and budgetary constrains in one model. Furthermore, Appendix 14 provides the averages of sample marginal effects for the models depicted in Appendix 13, including Model IV, that encompasses all the interacted variables with area.

Similar principle can be employed when interpreting interaction terms for the probit regression, as was the case in the interactions for hygiene behavior regression. However, rather than using coeffects, marginal effects need to be used when determinig the probability of the outcome – the change of a sanitation system.

Appendix 13, Model IV shows that only Barazas (not intreacted) and Earnings were a significant predictor when interacted with area, both in a negative direction.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Area	0.214***	0.204***	0.209***	0.230*	0.252	0.258*	0.399***
	(0.081, 0.347)	(0.071, 0.337)	(0.065, 0.352)	(-0.040, 0.499)	(-0.121, 0.625)	(-0.042, 0.557)	(0.110, 0.688)
Number of barazas		-0.048***	-0.050***	-0.046*	-0.050***	-0.050***	-0.049***
		(-0.080, -0.015)	(-0.083, -0.017)	(-0.094, -0.001)	(-0.083, -0.017)	(-0.083, -0.016)	(-0.082, -0.016)
Time spent on marketing (%)			0.022	0.022	0.028	0.024	0.027
			(-0.023, 0.067)	(-0.023, 0.067)	(-0.034, 0.089)	(-0.022, 0.069)	(-0.018, 0.073)
Main job			0.079	0.079	0.084	0.103	0.129
			(-0.126, 0.284)	(-0.126, 0.284)	(-0.125, 0.293)	(-0.139, 0.345)	(-0.087, 0.344)
Earnings			-0.009	-0.009	-0.010	-0.002	0.059
			(-0.193, 0.175)	(-0.193, 0.175)	(-0.194, 0.174)	(-0.190, 0.185)	(-0.146, 0.264)
Area:barazas (interaction)				-0.006			
				(-0.072, 0.060)			
Area:Time (interaction)					-0.011		
					(-0.097, 0.076)		
Area:Main (interaction)						-0.064	
						(-0.405, 0.278)	
Area:Earnings (interaction)							-0.257
							(-0.593, 0.080)
Constant	0.091*	0.262***	0.121	0.110	0.096	0.094	0.027
	(-0.003, 0.186)	(0.110, 0.413)	(-0.171, 0.414)	(-0.206, 0.427)	(-0.261, 0.453)	(-0.233, 0.421)	(-0.291, 0.345)
Observations	1,391	1,391	1,391	1,391	1,391	1,391	1,391
Log Likelihood	-941.579	-937.547	-936.932	-936.915	-936.902	-936.865	-935.809
Akaike Inf. Crit.	1,887.158	1,881.095	1,885.863	1,887.830	1,887.803	1,887.730	1,885.617

Note: *p**p***p<0.01

Appendix 12

	Dependent variable:			
	Change of the Sanitation System			
	(I)	(II)	(III)	(IV)
Area	0.214***	0.204***	0.209***	0.434
	(0.081, 0.347)	(0.071, 0.337)	(0.065, 0.352)	(-0.141, 1.008)
Number of barazas		-0.048***	-0.050***	-0.048**
		(-0.080, -0.015)	(-0.083, -0.017)	(-0.096, -0.0001)
Time spent on marketing (%)			0.022	0.034
			(-0.023, 0.067)	(-0.032, 0.100)
Main job			0.079	0.084
			(-0.126, 0.284)	(-0.175, 0.344)
Earnings			-0.009	0.081
			(-0.193, 0.175)	(-0.128, 0.289)
Area:Barazas (interaction)				-0.003
				(-0.069, 0.064)
Area:Time (interaction)				-0.018
				(-0.109, 0.074)
Area:Main job (interaction)				0.238
				(-0.272, 0.748)
Area:Earnings (interaction)				-0.441*
				(-0.930, 0.048)
Constant	0.091*	0.262***	0.121	0.014
	(-0.003, 0.186)	(0.110, 0.413)	(-0.171, 0.414)	(-0.418, 0.445)
Observations	1,391	1,391	1,391	1,391
Log Likelihood	-941.579	-937.547	-936.932	-935.201
Akaike Inf. Crit.	1,887.158	1,881.095	1,885.863	1,890.402

Note: *p**p***p<0.01

Appendix 13

	Regression I	Regression II	Regression III	Regression IV
(Intercept)	0.036*	0.101***	0.047	0.005
Area	0.083***	0.079***	0.08**	0.167
Number of barazas		-0.018***	-0.019**	-0.018*
Time spent on sanitation marketing (%)			0.008	0.013

Main Job			0.03	0.032
Earnings			-0.004	0.031
Area: Barazas (interaction)				-0.001
Area: Time (interaction)				-0.007
Area: Main job (interaction)				0.092
Area: Earnings (interaction)				-0.17*

Appendix 14