Reducing Child Mortality in the Last Mile: A Randomized Social Entrepreneurship Intervention in Uganda [*]

Martina Björkman Nyqvist, Andrea Guariso, Jakob Svensson, David Yanagizawa-Drott[†]

April 2017

Abstract

The delivery of basic health products and services remains abysmal in many parts of the world where child mortality is high. This paper shows the results from a large-scale randomized evaluation of a novel "social entrepreneurship" approach to health care delivery. In randomly selected villages a sales agent was locally recruited and incentivized to conduct home visits, educate households on essential health behaviors, provide medical advice and referrals, and sell preventive and curative health products. Results after three years show substantial health impact: under-5 child mortality was reduced by 27% at an estimated cost of $71 per life-year saved.

JEL classification: O12, I11, I12

Keywords: Child Mortality; Infant Mortality; Social Entrepreneurship; Community Health Worker; Uganda.

---

1. Introduction

Despite significant reductions in child and infant mortality over the last few decades, about one in twelve children in sub-Saharan Africa still die before his or her fifth birthday (UN, 2015). Many, if not most, of these deaths can be avoided through simple preventative care and through simple, low cost treatments delivered at home. This means that an effective response to reduce child deaths is not out of reach. While health outcomes can be tied to a host of factors, both on the demand and supply side, there is limited evidence on effective and scalable solutions to the problem (Dupas and Miguel, 2016). In this paper, we focus on the role of delivery on basic health services and products to poor communities at the very end of the supply chain. How to combat child mortality in this space is arguably of first order importance from a policy-making perspective. On the one hand, evidence shows that when household get access to very basic health products and services, free of charge, mortality is substantially reduced (Kumar et al, 2008; Baqui et al, 2008). On the other hand, even if such solutions are socially desirable in the long run, in the short run they imply severe feasibility challenges in countries with poor state capacity in general, and service delivery capacity in particular.

In this paper, more specifically, we study the impact of a novel "social entrepreneurship" approach – the *Living Goods model* – to community health delivery.[1] The model creates "Avon-like" networks of door-to-door mobile Community Health Promoters (CHP). The main activities of the CHPs resemble the standard activities of any community health worker: conduct home visits within their community, educate households on essential health behaviors, provide basic medical advice, and refer the more severe cases to the closest health center. On top of this, the CHPs make a modest income by selling a diverse basket of basic health goods, ranging from anti-malaria drugs to soap and fortified foods. CHPs purchase these products directly from the NGOs at wholesale price and earn

---

[1]The definition of social entrepreneurship is admittedly somewhat elusive, but typically entails a parallel approach of pecuniary (for-profit) and social goals. For a discussion, see Martin and Osberg (2007).

a margin on each product sold. Product prices are set by the NGOs in two dimensions: the price the CHPs are allowed to procure products at and their retail price. Products are effectively cross-subsidized to provide larger incentives on products that are believed to facilitate the greatest child mortality reductions, given existing market conditions. The underlying hypothesis is that these incentives on products, coupled with small financial incentives to encourage timely services[2], would not only move households down the demand curve but also motivate community health promoters to actively provide basic health services to mothers, newborn, and children. Thus, the CHPs operated as micro-entrepreneurs with financial incentives to meet household demand and improve child health.

The program was randomized across 214 rural villages spread across Uganda and was fully operational in all treatment clusters, with at least one CHP locally recruited to the program, by the beginning of 2011.[3] Our results show after three years, the intervention resulted in a substantial health impact: under-5 child mortality was reduced approximately by 27%, infant (i.e. under 1 year) mortality by 33% and neonatal (i.e. under 1 month) mortality by 28%.

While the evidence shows that households in treatment villages were more significantly likely to use products sold by the CHPs, such as insecticide treated bed nets and oral re-hydration salts for treating diarrhea – consistent with the effects on mortality –, we also provide suggestive evidence that the effects were not simply driven by access to cheaper, high quality, medicines and health products. In particular, the largest increases in the treatment relative the control group in terms of behavior were observed for follow-up visits and counseling. Households with a newborn baby were more than 70% more likely to have received a follow-up visit in the first week after birth, and households with

---

[2]The CHPs received small performance-based incentives to encourage registering of pregnant women and visits of newborns ($0.65 per registration/newborn visit).

[3]The rollout of the CHP program, including the trial clusters, was overseen by an advisory board including individuals with expertise in international public health and health service research as well as officials from the Uganda Ministry of Health.

a child under-five that fell sick with malaria or diarrhea were, respectively, 73% and 62% more likely to have received a follow-up visit. For households with infants that fell sick with malaria or diarrhea the increases were 109% and 105%, respectively. These results suggest that the efficacy of the program was not only a result of the pecuniary incentives to sell products, at least not directly, but also the provision of life-saving services. These findings add to the small but growing literature that investigates the role financial incentives can play in motivating individuals engaged in pro-social activities (Ashraf et al, 2014; Olken et al, 2014; Luo et al, 2015).

To facilitate policy making, we also assess the effectiveness of the program and perform a cost-benefit analysis taking into account fixed and variable costs. Our results indicate that the average cost per averted death was $4,237. This figure is about 40% of the cost per life saved that the Guttmacher Institute estimated could be achieved by expanding a range of health services known to be effective at saving lives (Perry and Zulliger, 2012). Taking into account that life expectancy in this context is vastly improved conditional on surviving the first few years of life, the estimated cost per life-year gained is equal to $71. This figure compares favorably to existing estimates, ranging from $82 per life-year gained in Kenya to $3,396 per life-year gained in Indonesia (Borghi et al, 2005; McPake et al, 2015).

The remainder of the paper is organized as follows. Details on the study setting, the research design, and the intervention are presented in section 2. Section 3 reports the main results. In section 4 we perform a robustness analysis. Section 5 presents the cost-effectiveness estimates and section 6 concludes.

2. Intervention and Empirical Design

2.1. The Program

In 2007 Living Goods, a US based NGO active in Uganda, in collaboration with BRAC Uganda began piloting a new community health delivery model intended to improve maternal, newborn and child health. Unlike volunteer-based community health worker programs, the community health promoter (CHP) program harnesses the power of franchised direct selling to provide CHPs with incentives to increase poor households' access to low-cost, high-impact health products and basic newborn and child health services. The CHP program was organized into geographically based branches, managed by branch managers and supervised by two NGOs (Living Goods and BRAC Uganda). Each CHP was assigned to a specific cluster, which in most cases corresponds to one village.

The CHP program is ongoing and by the end of the evaluation period in 2013 it was operating in 883 clusters (villages), organized in 29 branches, located in 23 districts, spread over all four regions of Uganda (see figure 1[4]). Thereafter the program continued to expand and by end of 2016 reached more than 5,500 clusters, organized in 143 branches, with a total population of over 4.4 million.

[FIGURE 1 ABOUT HERE]

The CHPs were selected through a competitive process among female community members aged 18 to 45 who applied for the position in each village and who possessed basic writing and math skills. Eligible candidates received 2 weeks of health and business training, covering preventing, diagnosing and treating childhood illness, recognizing danger signs for referral, healthy pregnancy and newborn care, and nutrition. At the

---

[4]Figure A.1 in Appendix provides a set of more detailed images by study district.

end of the training, a skills test was administered to determine who would become an active CHP. Selected CHPs also attended a one-day training each month to review and refresh key health and business topics.

The CHPs tasks were to conduct home visits to households with children under five years old, educate households on essential health behaviors, provide basic medical advice, referring the more severe cases to the closest health center, and to diagnose illness and sell preventive and curative health products. The CHPs were also instructed to visit newborns within the first 48 hours of life and to encourage pregnant women to deliver in a facility or with professional assistance.[5]

The product line the CHPs had at disposal included prevention goods (e.g. insecticide treated bednets, water purification tablets, and vitamins), curative treatments (e.g. oral re-hydration salts, zinc, and ACTs), as well as other health-related commodities (e.g. diapers, detergent, and hand soap) and durables with health benefits (e.g. improved cook stoves, solar lights, and water filters). As mentioned in the introduction, the broad product mix and the pricing strategy had three potential benefits: (i) driving up total sales and income for the CHPs; (ii) enabling the NGOs to cross-subsidize prices across products; (iii) motivating agents to be out visiting households regularly, thanks to the presence of high-velocity items in the product mix[6]. The products were sold by the CHP generally below prevailing market prices. The retail price was indeed determined by country management with a target of keeping prices for preventive and curative products on average 10% lower than the prevailing local market prices. The CHPs, in turn, purchased these products directly from Living Goods or BRAC branches at wholesale prices on average 30% below market prices and therefore earned an income on each product sold. Although the exact pricing strategy kept changing with the market conditions, CHPs

---

[5]Similar health services are in the public health literature often labeled as iCCM (integrated community case management) and MNCH (maternal, newborn, and child health) services.

[6]The business training received by the CHPs explicitly stressed the importance of building up a customer-base by providing free services like health education, referrals, and newborn visits. In addition, the CHPs received small performance-based incentives to encourage registering of pregnant women and visits of newborns ($0.65 per registration/newborn visit).

typically maintained higher margins on goods aimed to improve child health and reduce child mortality. Overall, the CHPs operated as micro-entrepreneurs with financial incentives to meet household demand and improve child health. The two NGOs also managed to keep their own margins at 10% at least, by buying the products largely directly from the producers, national importers, or national distributors.

2.2. Comparison with traditional community health workers and related literature

In many developing countries, as is the case in Uganda, the primary strategy to extend primary health care from facilities to under-served rural communities is community health workers (CHWs) (Singh and Sachs, 2013). In contrast with the *Living Goods model*, traditional CHW programs lack explicit monetary incentives (Bhutta et al, 2010; Christopher et al, 2011; Gilmore and McAuliffe, 2013; Haines et al, 2007; Perry and Zulliger, 2012). Specifically, community health work is often voluntary, but workers face competing opportunities such as paid-work or home production, that may lead them to devote less time to caregiving. Unsurprisingly, most of the evidence on the positive impact of CHW programs come from studies in settings with high quality supervision and support. Such a monitoring system may not be achievable in routine field situations. How to incorporate incentives to motivate CHWs in large-scale CHW programs, and the impact that will have, are open questions.

Systematic reviews of existing studies show that CHWs can be impactful in promoting positive health behavior and in providing basic curative and health services (Bhutta et al, 2010; Christopher et al, 2011; Gilmore and McAuliffe, 2013; Gogiaa and Sachdeva, 2010; Haines et al, 2007; Lewin et al, 2010; Naimoli et al, 2012; Okwundu et al, 2013; Perry and Zulliger, 2012). However, the findings from reviews of randomized controlled trials of CHW programs and CHW-led interventions are mixed (Lewin et al, 2010; Okwundu et al, 2013). Two proof-of-principle studies cited as evidence in the WHO and UNICEF

home-visits strategy statement documented large reductions in neonatal mortality (36-54%) (Kumar et al, 2008; Baqui et al, 2008). Four trials delivered in a program setting documented smaller (8-15%) – and in three out of four trials not statistically significant – impact (Darmstadt et al, 2010; Bhutta et al, 2011; Bhandari et al, 2012; Kirkwood et al, 2013). Two studies assessed the impact of community-based training of mothers, of which one focused on teaching mothers curative treatments of malaria, finding a 40% reduction in under-5 mortality (Kidane and Morrow, 2000), and one focused on teaching child care to expectant and postpartum women, finding instead no significant impact on neonatal and infant mortality (Sloan et al, 2008). Finally, two trials assessed the impact of Integrated Management of Childhood Illness program in Bangladesh (Arifeen et al, 2009) and Ethiopia (Amouzou et al, 2016), finding no significant effect on under-5 mortality. Our study adds to this literature by evaluating for the first time whether a social entrepreneurial approach to health care delivery can lead to significant improvements in children's health.

Our study also contributes to the literature on the effect of incentives for agents engaged in pro-social activities. Most literature in this area has focused on the education sector and on the impact of performance incentives for teachers (e.g. Lavy 2002; Glewwe et al, 2010; Muralidharan and Sundararaman, 2011; Duflo et al, 2012). Recently, however, four studies have looked into the impact of incentives on the delivery of health services. Within the context of rural China, Miller et al (2012) relied on a randomized trial to study the role of incentives for school principals to reduce anaemia among their students, finding a modest effect.[7] In Indonesia, Olken et al (2014) studied a program that links aid disbursements for health and education to the performance of health services, finding significant short-run improvements in health indicators. In Zambia, Ashraf et al (2014)

---

[7]In a new randomized controlled trial, performed in a similar setting, the authors find that large (but not small) incentives as well as larger unconditional grants succeed in reducing anemia substantially, with incentives being the most cost-effective solution. The authors also find that performance incentives and unrestricted grants work as substitutes, with the larger unconditional grants completely crowding-out the positive effect of incentives (Luo et al, 2015).

evaluated the effect of both financial and non-financial incentives on the performance of agents recruited by a public health organization, finding, among other things, that both types of rewards are effective when their relative value is high and that the effect is stronger for pro-socially motivated agents. Finally, within the context of government child care health workers in India, Singh (2015) showed that combining performance pay with information provision to mothers leads to a significant reduction in malnutrition, although individually the effects are negligible.

2.3. Study design and participants

The study was a parallel-group, stratified cluster randomized controlled trial, embedded in the roll-out of the CHP program. 214 clusters (rural villages) took part in the trial. The clusters were located in 12 geographical zones spread across Uganda (see figure 1). Within each zone, the clusters were randomly divided into a treatment group and a control group. In 11 zones out of 12 the randomization was balanced (1:1). In one zone and for operational purposes the randomization was unbalanced (2:1). At least one CHP was assigned to each cluster in the treatment group.[8] No CHP was assigned to the control clusters. All clusters were enumerated at baseline. There were on average 237 households per cluster at baseline.

The main objective of the trial was to assess the impact of having a CHP working in the cluster on improving children's health. The evaluation design and implementation was independent of program implementation.

The outcomes of interest were measured through a cross-sectional household survey administered between September and December 2013; approximately three years after the CHPs began operating in the treatment clusters. Before implementing the survey, each cluster was enumerated. A random computer-generated sequence was then used to se-

---

[8]One CHP was assigned to each cluster, with the exception of few large clusters, where two or three CHPs were assigned.

9

lect 40 households to be surveyed in each cluster (if less than 40 eligible households were available, all were sampled). The analysis was based on a final sample of 7,018 households, and their 11,563 under-5 children, that have lived in the same cluster throughout the trial. The final sample was slightly smaller than the cross-sectional household survey since households that migrated out from the baseline cluster were not included in the final analysis, nor were households that migrated into the trial clusters during the study period.

Sampled households were visited and asked for written informed consent to participate in the survey. Conditional on receiving the consent, an appointment was scheduled for the following day. The respondent was the female household head if available at the time of the interview or the primary female health care giver of the household. If neither could be found, or the household refused to participate, a replacement household was chosen (this happened in 7.2% of the cases, without any systematic difference between treatment and control clusters). Random back-checks were performed to ensure that all enumerators correctly followed the protocol.

The survey was implemented by Innovations for Poverty Action (IPA) Uganda and the survey teams were all composed by local staff with previous experience in data collection. Different survey teams operated in the different districts covered by the evaluation, to ensure that every staff member was familiar with local customs and spoke the local language. Data collectors were always masked to whether they were interviewing in an treatment or control cluster.

The trial was embedded in the rollout of the full CHP program (883 clusters) and there were no differences in program implementation between the treatment clusters (115 clusters) and the 768 clusters that were not part of the trial.

The CHPs were blinded to the trial status of the village they were assigned to avoid that the evaluation itself affected the CHPs behavior. As a consequence, no surveillance and monitoring system was put in place in the trial clusters and we did not track a pre-

10

determined set of households to avoid the CHP focusing their efforts on the households that were tracked at the expense of those who were not. Mortality rates were calculated based on cross-sectional household survey data collected at the end of the trial, using data from households that had resided in the same cluster throughout the trial. To ensure that these households were not systematically different in the two assignment groups, we tested for differential in- and out-migration during the trial period and checked for balance across assignment groups using pre-trial determined observable household characteristics, and pre-trial infant mortality rates, collected at the end of the trial period.

All households and especially households with children under-five were potential recipients of visits from the CHPs. While the CHPs were recommended to focus attention on providing services to households living within their cluster, they were not prevented from selling or providing advice also to households outside the cluster, including control clusters. Similarly, households living outside the treatment clusters could visit a CHP in an treatment cluster.

Households in both treatment and control clusters could benefit from primary health care services provided by other actors, including private clinics, public primary health dispensaries and village health teams (a government community health worker program).

2.4. Randomization

Figure 2 describes the trial profile. As the full CHP program was rolled out over time, the randomization of clusters was also phased in over time. We began in 2009 with a sample of 200 clusters (villages) in 10 geographic zones (8 districts). The clusters were stratified by zone and size (below or above 400 households) and, within each stratum, half of the clusters were assigned to the treatment group and half were assigned to the control group through a simple randomization procedure (computerized random numbers) generated by the researchers. In 2010, a year before the evaluation began; a decision was taken to only include clusters with less than 400 households at baseline as the design of the trial

was deemed less suitable for clusters where the CHPs only would be able to serve a small minority of the households. As a consequence, 10 strata with 94 clusters (47 treatment and 47 control) were deemed ineligible. 60 clusters organized in one new geographic zone were added in the end of 2010. Half of these 60 clusters were randomly assigned to the treatment group and half were assigned to the control group, following the same procedure adopted for the other zones. An additional zone was added in the beginning of 2011. For operational purposes, 1/3 of the 48 clusters in the final zone were randomly assigned to the control group and the remaining 2/3 of the clusters were allocated to the treatment group. The final sample for the trial thus consisted of 214 clusters (115 treatment and 99 control) in 12 zones (10 districts). Concerning the division between the two NGOs, 106 clusters (53 treatment and 53 control) were in BRAC-managed areas, while 108 were in LG-managed areas (62 treatment and 46 control). The program was fully operational in all treatment clusters in the beginning of 2011. Endline survey was conducted after three years, at the end of 2013.

[FIGURE 2 HERE]

2.5.   Outcomes

The pre-specified primary outcome was under-five mortality rate (U5MR). Secondary outcomes were infant mortality rate (IMR) and neonatal mortality rate (NMR). All mortality rates were calculated using the sample household survey data collected at the end of the trial. The household survey recorded detailed birth and death information for all children under five living in the households at the time of the survey as well as for all children that died under the age of five in the previous three years. For each child, we defined the number of month of exposure to the risk of death during the trial period, defined as the difference between the birth date of the child, or the start date of the trial (January 2011) if the child was born before that date, and the date that the child turned

12

five years if that occurred during the trial period, or the date of the endline household survey if the child was less than five years old at that time, or the date of the death of the child (see Figure A.2). Under-five mortality was then calculated as number of under-five deaths over the trial period per 1000 child-years of exposure to the risk of dying under the age of five. Infant mortality was calculated as number of deaths during the trial period arising within the first year of life per 1000 infant-years of exposure, with infant-years of exposure calculated in a similar way as the child-years of exposure to the risk of death. Neonatal mortality was calculated as number of deaths during the trial period within the first month of life per 1000 births.[9]

Additional secondary outcomes of interest were CHP interactions (to measure program coverage); follow-up visits; health knowledge and prevention; under-five morbidity (in self-reported malaria and diarrhea); treatment of under-five children for malaria and diarrhea; antenatal, delivery, and postnatal care.[10] Data on all secondary outcomes were collected in the endline sample household survey.

2.6. Sample size

The sample size was designed to detect a reduction in overall under-five mortality. In a community-based trial in 2009 with significant overlap in the regions covered to the CHP study, U5MR was 18 deaths per 1000 child-years with a coefficient of variation of the incidence rates of 0.32 (Björkman et al, 2017). On the basis of these data, and 120 child-years of observations in each cluster (i.e. three years and 40 child observations per

---

[9]International organizations such as UN and WHO typically express mortality in terms of deaths per 1000 live-births. Such organizations use data collected over long periods of time and rely on a life-table approach to compute mortality as a probability. Given that our evaluation lasted only for three years, the most appropriate approach in our case is to compute mortality as a ratio, following the steps described above, and to express it in terms of years of exposure. For completeness and in other to facilitate comparisons with other estimates, in the results section we will in any case also report results obtained using a life-table approach, although they should be considered less reliable.

[10]Pneumonia was initially included as a secondary outcome but due to changes in the regulatory environment, there was a delay in the authorization to include antibiotics among the list of health products provided by the CHPs and we could therefore not study pneumonia-related outcomes in the end.

year), a sample size of 214 clusters, of which 115 clusters are allocated to the treatment group and 99 clusters to the control group, would detect a 27% reduction in under-five mortality with 80% power at the two-sided 5% significance level.

3.   Results

3.1.   Balance at baseline

Table 1 reports balance tests using cluster-specific statistics before the program started. Results show that clusters were not statistically different between the treatment group and the control group in terms of size, household characteristics, and distance to main roads, electricity transmission lines, and health facilities.

[TABLE 1 HERE]

Household data was not collected at baseline by the research team.[11] We therefore use endline data to compute infant mortality for the two years preceding the intervention; i.e., in 2009 and 2010. Results in Table 2 show that there was no significant difference in infant mortality rates between the treatment and control prior to the intervention. IMR was 52.4 per 1000 child-years in the treatment group compared to 50.0 per 1000 child-years in the control group and the p-value on the difference was 0.83.

Panel B of Table 2 shows that pre-trial determined observable household characteristics, such as household size at the start of the trial and age and years of education of the household head, were not statistically different between the treatment group and the control group for the households used in the analysis; i.e., household that had remained in the same cluster throughout the trial and surveyed in 2013.

---

[11]Baseline data was collected for most BRAC clusters by BRAC itself. Although the analysis suggests a balanced sample, we do not include it in our discussion, due to its incomplete coverage and to the fact that data were collected by one of the implementing agencies.

3.2. Health outcomes

The primary outcome measure for the trial was child mortality. Child mortality links to the wide spectrum of child health specific services and health goods that the CHP provides to households with children under-5. Many of these services and goods, including health education, child preventative care, child curative care, and the medical drugs related to deworming, malaria, and diarrhea that were sold to affordable prices to the households, also have the potential to affect other health outcomes. We therefore also collected measures of height, weight and hemoglobin levels of all children under-5 living in the surveyed households.

We start by showing the findings on child mortality in table 3. The first three columns report the results using raw data: i.e. the number of under-five, infant (under 12 months), and neonatal (under 1-month) deaths per year. In order to assess the impact of the program, here as well as in the remaining of the paper, we compare mean outcomes after accounting for stratification. That is, we estimate

(1) $$Y_{ij} = \beta T_{ij} + b_j + \epsilon_{ij} \, ,$$

where Y_{ij} is the outcome of interest (e.g number of under-5 deaths over the study period) in village i, located in the geographical area associated to the NGO branch j. T is an indicator variable for villages assigned to the treatment groups, b_j are branch fixed effects, and ϵ is an error term.

The first three columns of the table show that the CHP program reduced the number of deaths in all three age categories. The number of under-5 deaths dropped by 28% (column i); the number of infant deaths dropped by 33% (column ii); the number neonatal deaths dropped by 27% (column iii). The raw data clearly shows that the CHP program had a

15

large impact on reducing mortality for young children in the treatment areas. However, the crude death numbers may not necessarily be due purely to a reduction in the risk of child death since cohort sizes may have been differentially affected by the intervention due to for example differences in fertility rates. Therefore, in column (iv)-(vi) of Table 3 we are estimating the mortality rate over the period of exposure; i.e. between January 2011 to December 2013. We follow the conventional approach used in epidemiology and define the under-five [infant] mortality rate as the number of under-five [infant] children that died during the period per 1000 child-years [infant-years] of exposure over the same time period. Conventionally, we define neonatal mortality as the number of neonatal deaths per 1000 live births.

The estimated rate ratio in column (iv); i.e. the incidence of child deaths in the treatment relative to the control group, implies that the risk of under-five deaths was reduced by 27% relative to the control group. The effect is of the same order of magnitude, and even more precisely estimated, using a linear model – a reduction of 5.95 deaths per 1000 child-years from a control group mean of 19.4 deaths per 1000. The reduction in infant mortality (column (v)) is even larger – a 33% reduction in the risk of infant deaths in the treatment versus the control group – and even more precisely estimated. Finally, the number of children dying before reaching one month (per 1,000 live births) is 24.1 in the treatment group compared to 33.4 in the control group and the difference – a reduction of 27.8% in neonatal mortality – is significant at the 5 percent level.

In columns (vii) and (viii) we also report under-5 and infant mortality expressed in terms of deaths per 1000 live births. In this case we estimate the probabilities of deaths by using a life-table approach. This is the methodology typically followed by international organizations such as UN and WHO, but it is more appropriate for data spanning longer time periods than the three years we are considering here. It however provides figures that are more easily comparable with those published by such organizations. WHO reports under-5 and infant mortality rates for Uganda in 2013 equal to 66.1 and 43.8, re-

spectively (UNICEF et al, 2014). Our estimates for control clusters are just slightly higher and equal to 68.4 and 49.7 deaths per 1000 live births, respectively. Columns (vii) and (viii) show that the intervention led to a reduction in treatment village of 19.9 deaths in mortality under-5 (28.9% reduction) and 17.2 deaths in infant mortality (34.7% reduction). In both cases the estimate is significant at 1%. These results appear very much in line with what we found above.

The main outcome of the CHP program was to reduce child mortality at all levels and the results presented in Table 3 clearly shows that the program indeed was successful in saving children's lives for all ages under five.

[TABLE 3 HERE]

Table 4 reports the results on weight, height, and hemoglobin levels of children under 5 years old. We measured weight using portable weighing scales and height using stadiometers. Hemoglobin levels were measured through a hemocue machine, which is a photometer that tests hemoglobin concentration using a single drop of blood taken from the child's finger.[12] The first four columns of table 4 report impact on height-for-age and weight-for-height, which are the standard measures used to identify stunting and wasting, respectively. In particular, according to WHO standards, a child is defined as moderately stunted (wasted) if the height-for-age (weight-for-height) value is below -2 standard deviations from the reference mean.[13] Results in columns (i) and (ii) suggest that the program led to an improvement in the height-for-age measure among children under-5, although the effects are not precisely estimated: the share of stunted children according to WHO standards in the treatment villages decreased by 1.9 percentage points,

---

[12]The samples were collected by selected staff members that followed a specific health training. The machine simply works by inserting a special glass (microcuvette) that contains the drop of blood just taken from the child's finger. The blood is analysed by the machine and the result is displayed in less than a minute. The microvuette containing the drop of blood was safely disposed immediately after the hemocule has revealed the hemoglobin level.

[13]Values are often expressed in terms of z-scores. The z-score records the anthropometric value as a number of standard deviations below or above the reference mean or median value.

which represents a 6.8 percent drop with respect to the prevalence in the control group, and the effect is significant at 10%. Results in columns (iii) and (iv) show that the program had no detectable effect on wasting. Finally, columns (v) and (vi) highlight a large and significant improvements in hemoglobin levels, with the share of anemic children (i.e. with hemoglobin level below 10g/dl) in the treatment group decreasing by 2.7 percentage points, or by 16% compared to average prevalence in the control group.

[TABLE 4 HERE]

3.3. Processes

The community health promoter program was intended to improve child health by providing basic curative and preventative health services. As we explained above, the CHPs tasks were to conduct home visits, educate households on essential health behaviors, provide basic medical advice, refer the more severe cases to the closes health center, visit newborn and to encourage pregnant women to seek antenatal care, as well as to sell preventative and curative health products. We have shown that the program had impact on the ultimate outcome – child mortality – and we now turn to assessing evidence on intermediate outcomes that relates to better child health.

Interaction and knowledge

We start by assessing the extent of interaction that households had with the community health promoters and whether the households in the treatment areas had improved health knowledge as part of the CHP's education purpose.

Column (i) of able 5 shows that almost 24% of the households in the treatment clusters have been visited by a CHP in the 30 days preceding the survey. While there was evidence of spillovers – 5.4% of the households in the control group have also been visited by a CHP

– households in the treatment group were more than 4 times as likely to have benefited from such a visit.

The CHP was supposed to educate households on basic health behaviors aimed to improve children's health. Column (i)-(vii) in Table 5 show that households in the treatment areas were indeed more informed about causes and treatments of diarrhea as well as about causes for malaria: compared to the control area, they were 11 percent more likely to know that diarrhea is transmitted by drinking untreated water; 16 percent more likely to know that zinc is effective in treating diarrhea; and 38 percent more likely to know that mosquito bites are the only cause of malaria. They were also 4.7 percentage points (control mean is 59.1%) more likely to have heard of food with added vitamins or nutrients. Knowledge about bednets and the importance of professional assistance when giving birth did not differ between control and treatment groups, although in these cases there was limited room for improvement, as even in control villages more than 99% of respondents displayed such knowledge.

In the last column of table 5 we report average standardized effects of the health knowledge outcomes; i.e. we estimate a seemingly unrelated regression system

$$(2) \qquad\qquad Y = [I_n \otimes T]\beta + \mu \, ,$$

where Y is a vector of n related health knowledge outcomes, I_n is a n by n identity matrix, and T is a vector of assignment to treatment group indicators. We derive an average standardized effect, $\tilde{\beta} = \frac{1}{n} \sum_{n=1}^{N} \frac{\hat{\beta}_n}{\hat{\sigma}_n}$ where $\hat{\beta}_n$ is the point estimate on the treatment indicator in the nth outcome regression and $\hat{\sigma}_n$ is the standard deviation of the control group for outcome n (see Kling et al., 2004; Duflo et al., 2008). The average standardized effect is positive and highly significant, confirming that overall the intervention improved households health knowledge.

[TABLE 5 HERE]

19

Health behavior and morbidity

Table 6 depicts the results on the households' preventative health actions as well as morbidity. In principle, these self-reported outcomes on the individual's health behavior could be difficult to interpret because of the potential recall and social desirability biases in self-reported data (Strauss and Thomas, 1998; Powers et al, 2008). However, in this context these results should be viewed as complements to the more robust health outcomes of mortality and anthropometric data. The findings in Table 6 reveal that households in the treatment group were 3.8 percentage points (control mean is 77.4%) more likely to have treated their water before use and their children were 13 percent more likely to have slept under an insecticide-treated bednet. These outcomes can be linked to the presence of the CHP, who is incentivized to promote and sell both of the health products associated to this preventative behavior – water purification tablets and insecticide treated bednets – to households with young children.

Self-reported morbidity in malaria and diarrhea did not differ between control and treatment groups. Conditional on falling sick with malaria children in the treatment group were as likely as children in the control group to have received treatment with ACTs for at least 3 days. However, conditional on falling sick with diarrhea children in the treatment group were 16.2 percent more likely to have received treatment with ORS/Zinc.

The average standardized effect reported in column (viii) indicates that overall the intervention led to significant improvements in household health behavior.

[TABLE 6 HERE]

Health visits

One of the important tasks of the CHP was to build up a stable customer-base by providing health education, referrals, and newborn visits. Follow-up visits and revisits to sick

20

households would arguably create trust between the health worker and the household. Table 8 looks at the occurrence of follow-ups among households. Households in the treatment area with a newborn baby were 71% more likely to have received a follow-up visit in the first week after birth compared to households in the control areas. Similarly, households with a child under-five in the trial areas that fell sick with malaria or diarrhea were, respectively, 73% and 62% and more likely to have received a follow-up visit compared to households with sick children in the control areas. For households with *infants* that fell sick with malaria or diarrhea the increases were 7.3 and 8.1 percentage points, respectively, which means an increase of more than 100% compared to households with sick infants in the control areas. These are all cases that the CHP is particularly trained and incentivized to focus on specifically newborn visits and children sick with malaria and diarrhea and consequently, we also see an increase in follow-up visits for the youngest children. The average standardized effect in column (vi) confirm that the program led to a large and significant increase in the likelihood of follow-up visits.

When looking at counseling, we also see that a significantly higher share (about 10%, significant at the 5 percent level) of women in the treatment group had been advised to give birth with professional assistance, although the share that gave birth in a health facility and the share of the currently pregnant women that had received at least some antenatal care did not differ between control and treatment groups.

[TABLE 7 HERE]

4. Robustness

Our study has some limitations. First, the choice not to have surveillance or monitoring systems in place in the study villages implied that we had to rely on retrospective recall information. We used standardized data collection methods, and any potential recall

lapses were expected to affect the treatment and control groups equally and thus lead to an attenuation bias that would lead us to estimate a lower bound on the impact of the CHP program on the outcome variables, including child mortality estimates.

Second, as we use the end of trial sample survey to define baseline residence and thus the core sample for the analysis, selective out-migration by assignment groups could have caused some confounding bias in our main estimates. In table 8 we test whether there was selective in- and out-migration using enumeration data at baseline and endline combined with data from the household survey.

[TABLE 8 HERE]

At baseline 50,617 households were residing in the trial cluster, 4,132 of whom were estimated to have migrated out from the baseline cluster by the end of the trial. The average rate of out-migration per cluster was 7.1% and was not statistically different between the treatment group and control group (p=0.991). An estimated 7,962 households moved into the trial clusters during the study period. The average rate of in-migration per cluster was 15.3% and was not statistically different between the treatment group and control group (p=0.478). The share of sampled households that has moved into the cluster during the trial period, out of the total number of sampled households, was not statistically different between the treatment group and control group (p=0.614). Hence, measured in- and out-migration into the study clusters were similar across assignment groups. Moreover, as showed in table 2, baseline household characteristics of the sampled households that had lived in the same cluster for the whole study period were not statistically different between the treatment group and the control group. Overall, these results indicate that the results are unlikely to be biased by migration patterns.

Third, the possibility of contamination is plausible because the study clusters, within each zone, were geographically close. Analysis of behavioral data (interaction with CHP in the control sites) also suggested that some, although low, contamination occurred, most

likely causing us to estimate a lower bound on the impact of the CHP program on child mortality.

Finally, another potential concern when studying this program is that charging for preventive and curative products, even when prices are low, will disproportionately benefit the less-poor households. Table A.1, however, suggests similar impact across the household wealth distribution.

5.   Cost effectiveness analysis

In the previous sections we have shown that the incentivized community health worker program resulted in a large reduction in child mortality for all ages under five. This particular program was aimed to make the community health promoters self-sustainable, by allowing them to become a micro-entrepreneur selling basic health products at prices below the market price and at the same time earning a margin while also providing preventative and basic health services. It is therefore important to investigate the cost-effectiveness of the program.

Community health worker programs have typically been promoted as an effective and relatively cheap way to deliver health services in low- and middle-income countries (Dahn et al, 2015). Remarkably, there is a dearth of information on the actual cost-effectiveness of such programs (Bhutta et al, 2010; Vaughan et al, 2015). In many cases estimates are computed using the Lives Saved Tool (LiST) (McPake et al, 2015). The LiST relies on an empirical model to estimate, among other things, how a projected change in inputs – typically the coverage rates of some health interventions – translates into a reduction in child mortality. The model uses national demographic data to produce estimates of lives saved in a national population, for given assumptions on the coverage of the intervention and its impact. With the data at our disposal, we can compute the cost effectiveness of the CHP program by taking into account the *actual* impact of the program

observed in our study.

In estimating cost effectiveness, we take expenditure figures from the budgets provided by Living Goods, one of the two NGOs that managed the CHP program.[14] Over the study period the 46 Living Goods control clusters hosted an estimated 5,339 children under five, 194 of which died before reaching the age of 5. By applying this mortality rate to the 8,306 children located in the 62 Living Goods treatment clusters, we would have expected 302 deaths under-5 in this group of villages in the absence of the intervention. The program led to a 31% reduction in child mortality in the Living Goods areas[15], i.e. 93.5 averted deaths in the treatment clusters over the three year period of the RCT. This means an average of 31.2 deaths averted per years. There were overall 95 CHPs operating in the Living Goods treatment villages during the study period. This translates in an average of 0.328 deaths averted per CHP per year. By 2013 Living Goods had a total of 466 community health promoters operating across Uganda, which implies a total of 153 deaths averted within a year. The total yearly cost of running the CHP program in Uganda in 2013 was $647,841. This figure includes: all country-level expenses (57%), expenses related to branch offices and transportation (22%), expenses for marketing and promotion (10%), and training (6%). Importantly, this figure does not include the original cost of health-related goods bought by Living Goods ($247,904), as these were subsequently sold to the CHPs. The profits made by Living Goods on these sales ($73,356) have been deducted from the cost figure. Based on these figures, the estimated cost per averted death under-five was $4,237.

---

[14]Although, as explained, the program was run by two different NGOs – Living Goods and BRAC – for the cost effectiveness calculation we only rely on the detailed budget provided by Living Goods. The reason is that Living Goods only focuses on implementing the CHP program, while BRAC has a vast portfolio of activities, including microfinance and agriculture extension programs. Given the presence of synergies across programs, it is difficult to isolate the portion of costs to be attributed to the CHP program. In any case, due to such synergies, the costs born by BRAC would be significantly lower than those born by Living Goods: a (self-reported) estimate of the net cost per capita served by the CHPs, is $2.07 for Living Goods and $0.55 and BRAC. These lower costs would be less representative for another organization or institution that wanted to replicate the CHP model. For all these reasons, we rely on the Living Goods budget and for simplicity, we focus on the 2013 budget, i.e. the most recent budget to the study period.

[15]The estimated mortality reduction was 31% (pvalue=0.06) in the Living Goods sample and 23% (pvalue=0.08) in the BRAC sample, although the difference is not statistically significant.

This figure is about 35% of the cost per life saved that would be achieved by expanding a range of health services known to be effective at saving lives ($12,000), as estimated by the Guttmacher Institute based on a range of available cost-effectiveness estimates (Perry and Zulliger, 2012). The Guttmacher Institute does not explicitly refer to community health workers, but the activity of the CHWs appears essential in achieving the proposed expansion in service provision. Overall, however, it is difficult to put the $4,237 estimate in perspective, given the lack of estimates of CHW programs based on rigorous evaluations. One way of looking at cost effectiveness, is to consider how much each life saved would have contributed in economic activity over his or her lifetime. Dahn et al (2015), for instance, estimate that on average, across SSA, a child under-5 will contributes approximately $64,645 in economic activity over his or her lifetime.[16] Based on this estimate, investment in the CHP program can result in an economic return of more than 15:1.

An alternative approach is to estimate the cost of the intervention in terms of life-year gained and compare it with the estimates obtained through the LiST. According to WHO estimates, life expectancy at age 5 in Uganda in 2013 was 59.5. The 153 deaths under 5 averted by the program would therefore translate in 9,104 life years saved each year. Overall, this means a cost of $71.2 per life-year saved.[17] This compares favorably to most estimates based on LiST for a range of community health workers programs and that have been found to vary from $82 per life-year gained in Kenya to $3,396 per life-year gained in Indonesia (Borghi et al, 2005; McPake et al, 2015).

Overall, we consider the $4,237 figure to be a conservative estimate of the cost per life saved for the CHP program. First, the evaluation took place when the program was still relatively new. As the program expands, with more villages being reached, and more

---

[16]The estimate is based on the following assumptions: i) GDP per-capita $1,738; ii) expected GDP per-capita growth of 2.5% per annum; iii) child will enter the workforce at age 18 and exit the workforce at age 56; iv) discount rate of 5% to calculate the net present value of the future cash-flows from projected lifetime earnings.

[17]For simplicity in all these calculations we have abstracted away from discount factor issues and uncertainty of point estimates.

CHPs being recruited, cost effectiveness is expected to improve. For instance, based on expansion (and budget) forecasts, Living Goods self-reported estimates on the cost per capita served by its CHPs are expected to drop from $2.07 in 2014 to $1.21 in 2018.

Secondly, it should be kept in mind that cost per life saved or per life-year gained are commonly used metrics to assess the cost effectiveness of a health intervention, but there are many other dimensions that could be taken into account whenever considering a community-based intervention. Some of the main advantages of the CHP program are indeed impossible to monetize and could not be included in the cost-effectiveness evaluation presented above. For instance, the fact that the program makes many more households able to rely on primary health-care in their own village, rather than traveling kilometers away to reach a health facility, is likely to bring a range of benefits to the society: it lowers the time and transportation costs for the patient, while also reducing the workload of the health facilities, freeing up resources that can be better used for other more urgent cases. Moreover, the presence of financial incentives provides all CHPs and their families with an additional source of income.

6. Conclusion

We estimate that the CHP program in Uganda reduced under-five mortality rate by 27%, infant mortality rate by 33%, and neonatal mortality rate by 27% after 3 years. These effects are supported by changes in health knowledge, preventive behavior, case management of malaria and diarrhea, and home visits. We also estimate the cost-effectiveness of the program and find the estimated cost per averted death under-five during the study period to be $4,237. To the best of our knowledge, this is the first assessment of the cost effectiveness of a community health worker program, based on a rigorous RCT approach.

While a growing body of evidence has identified effective interventions that can be delivered by community health workers, a key consideration for the success and sustain-

ability of such programs is how high-quality performance by community workers can be achieved and maintained. This study is the first impact evaluation of a community health delivery intervention based on an incentivized approach. Unlike previous studies that have primarily focused on the impact of specific interventions that could be delivered effectively in a community setting, our focus is on how to ensure that community health workers successfully implement a set of interventions proven to be effective if delivered and the impact that may have on child health.

In the CHP program, community health workers operated as micro-entrepreneurs earning an income on the sale of preventive and curative products. A concern with such a scheme is that it may encourage overuse of medications and inappropriate treatment at the expense of prevention and referrals. On the other hand, the provision of free services like health education and follow-up visits was viewed as strategy to build up a loyal customer base. More generally whether extrinsic incentives in some domains have positive or negative impacts on intrinsic motivation in other domains is an empirical question. The data does not suggest that the program only had an impact on incentivized services, with evidence of increases in the promotion of healthy behavior and changed health beliefs. While there was a large increase in visits of newborns, for which the CHPs received a small incentive payment, there were also large increases in follow-up visits of children sick in malaria and diarrhea, for which no direct incentives were attached.

Our analysis has shown no significant impact on malaria prevalence and treatment. However, similar treatment pattern does not necessarily imply similar quality of treatment. Among other things, the CHPs sell authentic ACT drugs. In the private market there is growing evidence that the market for antimalarial medicines is plagued by counterfeit and substandard (fake) products, with recent estimates suggesting that as much as a third of the antimalarial drugs sold contain too little or no active pharmaceutical ingredients (Nayyar et al, 2012). Uganda is no exception: a smaller study conducted in the same research areas one year into the program estimated that 37 percent of the retail

27

outlets were selling substandard antimalarial drugs (Björkman et al, 2014). Poor quality is not specific to ACTs but is a generic problem in the largely unregulated market for preventive and curative health products in many developing countries. The CHPs market share for ACT drugs and ORS were 11.3% and 14.1% respectively. Under the assumption that every third dose of ACT treatment sold in the private market is fake and that authentic drugs are provided in the public sector (about 40% of the market share), children in the treatment group are 19% less likely to be treated with a fake ACT medicine.

It is possible that the CHP program affected child mortality not only through the provision of curative and preventative maternal, child, and newborn services, but also through the subsidized sale of other health-related commodities and durables (e.g. hand soap, improved cook stoves, fortified food, and water filters). The broad product mix, with high-velocity items like soap and fortified foods, and low-velocity but high returns per sold unit items like improved cook stoves, was deemed crucial to motivate agents to be out visiting households regularly and for driving up total sales and income for the CHPs.

With the accumulated know-how we have today few would question the potential of community health care provision. How to best ensure that CHW deliver timely and appropriate services is, however, largely an open question and motivates the continued search for innovative approaches. The CHP program we studied here harnesses the power of franchised direct selling (business-in-a-bag) to provide community health providers with incentives to increase access to low-cost, high-impact health products and basic newborn and child health services. As of the end of 2016, the program was active in over 5,500 clusters with a total population of over 4.4 million and the scale-up is continuing and by end of 2018 it is estimated to reach over 5.3 million individuals in more than 6,700 clusters. The impact of the CHP program was conditional on existing facility based professional health care as availability of referral services is a crucial component to the program. Thus the findings should encourage government and non-government organi-

zations to continue improving their facility based care, but also points to the importance of integrating the program into the existing health service provision strategy. The process of integrating the CHP program we have evaluated here into the overall health care system is currently underway.

References

[1] AMOUZOU, A., E. HAZEL, B. SHAW, N.P. MILLER, M. TAFESSE, *et al*. Effects of the integrated Community Case Management of Childhood Illness Strategy on Child Mortality in Ethiopia: A Cluster Randomized Trial. *American Journal of Tropical Medicine and Hygiene* (2016), 94(3), 596-604.

[2] ARIFEEN, S.E., D.M. HOQUE, T. AKTER, M. RAHMAN, M.E. HOQUE *et al*. Effect of the Integrated Management of Childhood Illness strategy on childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. *Lancet* (2009), 374(9687): 393-403.

[3] ASHRAF, N., O. BANDIERA, AND J. B. KELSEY. No margin, no mission? A field experiment on incentives for public service delivery. *Journal of Public Economics* (2014) 120, 1-17.

[4] BAQUI, A.H., S. EL-ARIFEEN, G.L. DARMSTADT, S. AHMED, E.K. WILLIAMS *et al*. Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *Lancet* (2008), 371, 1936-1944.

[5] BHANDARI, N., S. MAZUMDER, S. TANEJA, H. SOMMERFELT, AND T.A. STRAND. Effect of implementation of Integrated Management of Neonatal and Childhood Illness (IMNCI) programme on neonatal and infant mortality: cluster randomised controlled trial. *BMJ* (2012), 344, e1634.

[6] BHUTTA, Z.A., S.L. ZOHRA, G. PARIYO, L. HUICHO. *Global Experience of Community Health Workers for Delivery of Health Related Millennium Development Goals: A Systematic Review, Country Case Studies, and Recommendations for Integration into National Health Systems*. WHO Geneva (2010).

[7] BHUTTA, Z.A., S. SOOFI, S. COUSENS, S. MOHAMMAD, Z.A. MEMON ZA *et al*. Improvement of perinatal and newborn care in rural Pakistan through community-based strategies: a cluster-randomised effectiveness trial. *Lancet* (2011), 377, 403-412.

[8] BJÖRKMAN, M., D. DE WALQUE, AND J. SVENSSON. Experimental Evidence of the Long-Run Impact of Community Based Monitoring. *American Economic Journal: Applied Economics* (2017), 9(1), 33-69.

[9] BJÖRKMAN. M., J. SVENSSON, AND D. YANAGIZAWA-DROTT. The Market for (Fake) Antimalarial Medicine: Evidence from Uganda. *Mimeo* (2014).

[10] BORGHI, J., B. THAPA, D. OSRIN, S. JAN, J. MORRISON *et al*. Economic assessment of a women's group intervention to improve birth outcomes in rural Nepal. *Lancet* (2005), 366(9500), 1882-1884.

[11] CHRISTOPHER, J.B., A.L. MAY, S. LEWIN, AND D.A. ROSS. Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa. *Human Resources for Health* (2011), 9-27.

[12] DAHN, B., A.T. WOLDEMARIAM, H. PERRY, A. MAEDA, D. VON GLAHN *et al*. *Strengthening Primary Health Care through Community Health Workers: Investment Case and Financing Recommendations*. Joint Report by John Hopkins University, the Office of the UN Special Envoy for the health-MDGs, the World Bank, Partners in Health, Last Mile Health, the Clinton Foundation, ALMA and the governments of Ethiopia and Liberia (2015).

[13] DARMSTADT, G.L., Y. CHOI, S.E. ARIFEEN, S. BARI, S.M. RAHMAN *et al*. Evaluation of a cluster-randomized controlled trial of a package of community-based maternal and newborn interventions in Mirzapur, Bangladesh. *PLoS One* (2010), 5, e9696.

[14] DUFLO, E., R. GLENNERSTER, AND M. KREMER. Using Randomization In Development Economics Research: A Toolkit. In T. Schultz and John Strauss (eds.) *Handbook of Development Economics* (2008), Vol. 4, Amsterdam and New York: North Holland.

[15] DUFLO, E., R. HANNA, AND S. RYAN. Incentives work: Getting teachers to come to school. *The American Economic Review* (2012), 102(4), 1241-1278.

[16] DUPAS, P., AND E. MIGUEL. *Impacts and Determinants of Health Levels in Low-Income Countries*. Handbook of Field Experiments, forthcoming, (eds.) E. Duflo & A. Banerjee.

[17] GILMORE, B., AND E. MCAULIFFE. Effectiveness of community health workers delivering preventive interventions for maternal and child health in low- and middle-income countries: a systematic review. *BMC Public Health* (2013), 13, 847.

[18] GLEWWE, P., N. ILIAS, AND M. KREMER. Teacher incentives. *American Economic Journal: Applied Economics* (2010), 2(3), 205-227.

[19] GOGIAA, S., AND H.S. SACHDEVA. Home visits by community health workers to prevent neonatal deaths in developing countries: a systematic review. *Bull World Health Organ* (2010), 88, 658-666.

[20] HAINES, A., D. SANDERS, U. LEHMANN, A.K. ROWE, J.E. LAWN *et al*. Achieving child survival goals: potential contribution of community health workers. *Lancet* (2007), 369, 2121-2131.

[21] KIDANE, G., AND R.H. MORROW. Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial. *Lancet* (2000), 12, 356(9229), 550-555.

[22] KLING, J.R., J.B. LIEBMAN, L.F. KATZ, AND L. SANBONMATSU. Moving To Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment. *KSG Working Paper* No. RWP04-035 (2004).

[23] KIRKWOOD, B.R., A. MANU, A.H. TEN ASBROEK, S. SOREMEKUN, B. WEOBONG *et al*. Effect of the Newhints home-visits intervention on neonatal mortality rate and care practices in Ghana: a cluster randomised controlled trial. *Lancet* (2013), 22, 381(9884), 2184-2192.

[24] KUMAR, V., S. MOHANTY, A. KUMAR, R.P. MISRA, M. SANTOSHAM *et al*. Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. *Lancet* (2008), 372, 1151-1162.

[25] LAVY, V. Evaluating the effect of teachers? Group performance incentives on pupil achievement. *Journal of Political Economy*, 2002, 110(6), 1286-1317.

[26] LEWIN, S., S. MUNABI-BABIGUMIRA, C. GLENTON, K. DANIELS, X. BOSCH-CAPBLANCH *et al*. Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases (Review). *The Cochrane Library* (2010), 3.

[27] LUO, R., S. ROZELLE, S. SYLVIA, AND M. VERA-HERNÀNDEZ. Can Bureaucrats Really be Paid Like CEOs? School Administrator Incentives for Anemia Reduction in Rural China. *NBER Working Paper 21302* (2015).

[28] MARTIN, R.L., AND S. OSBEG. *Social Entrepreneurship: The Case for Definition*. Stanford Social Innovation Review (2007) (Available at `http://ssir.org/articles/entry/social_entrepreneurship_the_case_for_definition`).

[29] MCPAKE, B., I. EDOKA, S. WITTER, K. KIELMANN, M. TAEGTMEYER *et al*. Cost-effectiveness of community-based practitioner programmes in Ethiopia, Indonesia and Kenya. *Bulletin of the World Health Organization* (2015), 93, 631-639.

[30] MILLER, G., R. LUO, L. ZHANG, S. SYLVIA, Y. SHI *et al*. Effectiveness of provider incentives for anaemia reduction in rural china: A cluster randomised trial. *BMJ* (2012), 345, 1-10.

[31] MURALIDHARAN, K., AND V. SUNDARARAMAN. Teacher performance pay: Experimental evidence from india. *Journal of Political Economy* (2011), 119(1), 39-77.

[32] NAIMOLI, J.F., D.E. FRYMUS, E.E. QUAIN, AND E.L. ROSEMAN. *Community and Formal Health System Support for Enhanced Community Health Worker Performance: A U.S. Government Evidence Summit*. USAID Final Report (2012).

[33] NAYYAR, G., J.G. BREMAN, P.N. NEWTON, AND J. HERRINGTON. Poor-quality Antimalarial Drugs in Southeast Asia and Sub-Saharan Africa. *Lancet Infectious Diseases* (2012), 12(6), 488-496.

[34] OKWUNDU, C.I., S. NAGPAL, A. MUSEKIWA, AND D. SINCLAIR. Home- or community-based programmes for treating malaria (Review). *The Cochrane Library* (2013), 5.

[35] OLKEN, B. A., J. ONISHI, AND S. WONG. Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia. *American Economic Journal: Applied Economics* (2014), 6(4), 1-34.

[36] PERRY, H., AND R. ZULLIGER. *How Effective Are Community Health Workers? An Overview of Current Evidence with Recommendations for Strengthening Community Health Worker Programs to Accelerate Progress in Achieving the Health-Related Millennium Development Goals.* Johns Hopkins Bloomberg School of Public Health (2012).

[37] SINGH, P., AND J.D. SACHS. 1 million community health workers in sub-Saharan Africa by 2015. *Lancet* (2013), 382, 363-365.

[38] SLOAN, N.L., S. AHMED, S.N. MITRA, N. CHOUDHURY, M. CHOWDHURY, U. ROB, AND B. WINIKOFF. Community-based kangaroo mother care to prevent neonatal and infant mortality: a randomized, controlled cluster trial. *Pediatrics* (2008), 121(5).

[39] UNICEF, WHO, THE WORLD BANK, UN POPULATION DIVISION. *Levels and Trends in Child Mortality*. Report (2014).

[40] VAUGHAN, K., M.C. KOK, S. WITTER, AND M. DIELEMAN. Costs and cost-effectiveness of community health workers: evidence from a literature review. *Human Resources for Health* (2015), 13, 71.

FIGURES & TABLES

Figure 1: Map of Districts and Distribution of Clusters
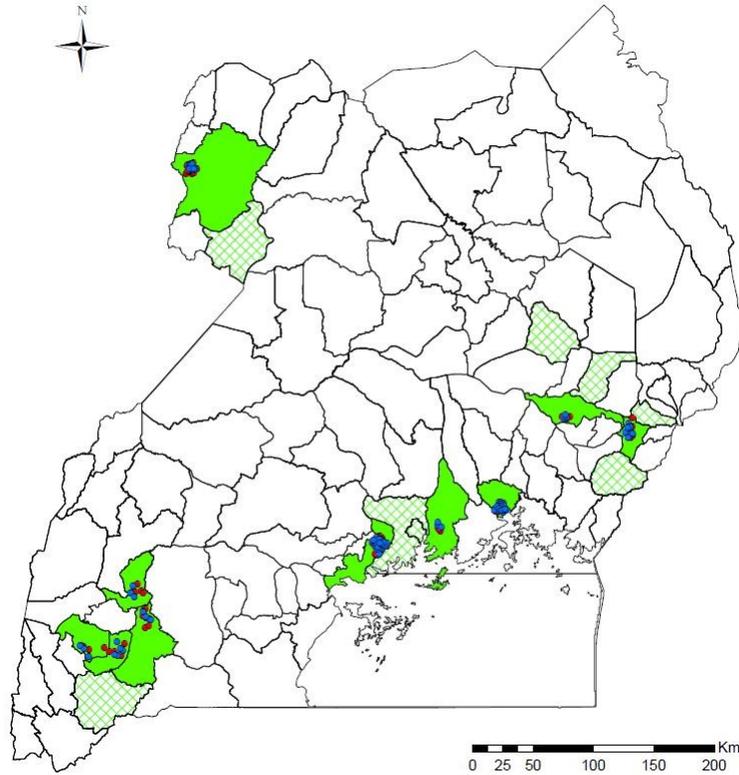
Notes: Green fully-colored areas indicate districts that were part of the study, while cross-hatched areas indicate districts excluded from the study, but in which the program was also implemented. Red and blue dots indicate respectively control and intervention villages included in the study. Figure A.1 in Appendix provides a set of more detailed images by study district.

Enrolment

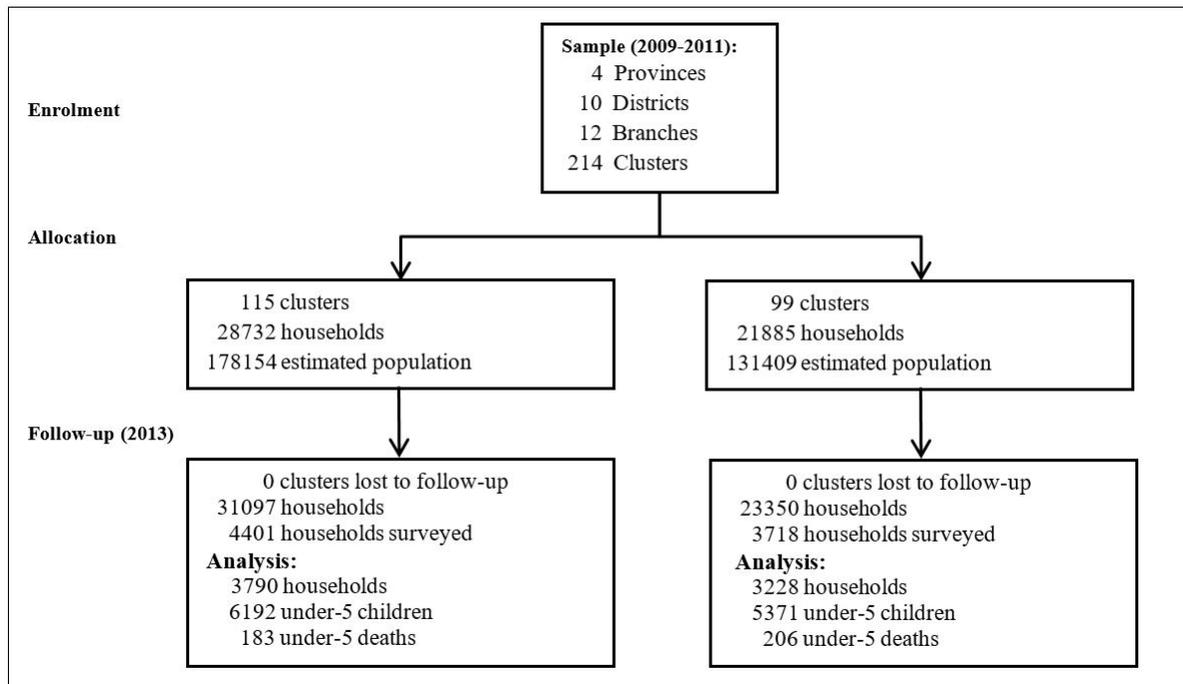**Sample (2009-2011):**
  4 Provinces
 10 Districts
 12 Branches
214 Clusters

Allocation

115 clusters
28732 households
178154 estimated population

99 clusters
21885 households
131409 estimated population

Follow-up (2013)

0 clusters lost to follow-up
31097 households
4401 households surveyed
Analysis:
  3790 households
  6192 under-5 children
  183 under-5 deaths

0 clusters lost to follow-up
23350 households
3718 households surveyed
Analysis:
  3228 households
  5371 under-5 children
  206 under-5 deaths

Table 1: Baseline Characteristics

Variables	Treatment Group	Control Group	p-value
Number of clusters	115	99	
Households per cluster	250 (113)	221 (107)	0.226
Households with under-5 children per cluster	86 (47)	78 (46)	0.665
Distance to main road	5.6 (11.6)	6.8 (12.7)	0.126
Distance to electricity transmission line	1.8 (1.5)	1.8 (1.5)	0.707
Distance to health center	1.4 (1.1)	1.7 (1.2)	0.256
Number of health centers within 5 km	8.3 (5.0)	7.3 (5.2)	0.459
Distance to hospital	10.4 (8.5)	11.1 (8.5)	0.916

Notes: Cells report mean (SD) across clusters included in the treatment or control group. A variety of sources were consulted to generate the original dataset, including documents and maps from national utilities, regional power pools, and the World Bank. Information on households and households with under-5 children per cluster was collected from the enumeration of trial villages at baseline. Data for medium and high voltage electricity transmission lines was obtained from the Africa electricity transmission network (AICD) study. Health Centers takes into account facilities from HCIII (i.e. parish-level health centers, roughly one per 5,000 people) and above. Hospitals refer only to district/national hospitals (roughly one per 500,000 people). Distance measures are all expressed in kilometers.

Table 2: Baseline Characteristics of Households not Lost to Follow-up and Surveyed at Endline

Variables	Treatment Group	Control Group	p-value
A. Infant mortality			
Years of exposure to risk of death under 1 year	1927	1743	
Deaths under 1 year	101	87	
Mortality rate per 1000 years of exposure	52.4	50.0	0.830
B. Households			
Number of household	3787	3217	
Household size	5.2 (2.3)	5.3 (2.3)	0.518
Age household head	36.4 (12.1)	36.7 (12.4)	0.641
Years of education household head	8.0 (0.4)	8.0 (0.2)	0.320

Notes: Cells report mean (SD) from endline sample household survey data for household that have remained in the cluster throughout the trial, with values scaled back to baseline period.

Table 3: Program Impact on Child Mortality

Dependent variable	Number of deaths			Mortality per 1000 yrs of exposure		Mortality per 1000 live births		
	Under-5 deaths (i)	Infant deaths (ii)	Neonatal deaths (iii)	Under-5 mortality (iv)	Infant mortality (v)	Neonatal mortality (vi)	Under-5 mortality (vii)	Infant mortality (viii)
Program impact	-0.58**	-0.54***	-0.29*	-5.95***	-18.87***	-9.27**	-19.86***	-17.26***
	(0.23)	(0.19)	(0.15)	(2.06)	(5.94)	(4.62)	(7.23)	(5.35)
Rate ratio				0.73**	0.67***	0.73**		
				(0.09)	(0.09)	(0.11)		
Mean control	2.08	1.62	1.07	19.4	52.7	33.36	68.4	49.7
Observations	214	214	214	214	214	214	214	214
R-squared	0.148	0.163	0.162	0.145	0.155	0.135	0.117	0.153

Notes: *Program impact* measures the coefficient on the assignment to treatment indicator, from a standard OLS regression. Dependent variables: (i) number of under-5 deaths; (ii) number of infant deaths; (iii) number of neonatal deaths; (iv) number of under-five deaths per 1000 child-years of exposure to the risk of deaths; (v) number of infant deaths per 1000 child-years of infant exposure to the risk of deaths; (vi) Number of neonatal deaths per 1000 births; (vii) Number of under-five deaths per 1000 births; (viii) Number of infant deaths per 1,000 births. Rate ratios in specifications (iv)-(vi) are derived from a Poisson model, with branch fixed effects and standard errors clustered by village; the number of observations are 11,342 (iv), 8,808 (v), and 6,499 (vi). Branch fixed effects are included in every regression. There are 12 branches in the sample. Robust standard errors in parentheses. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 4: Program Impact on Child Weight, Height, and Hemoglobin Levels

Dependent variable	Height-for-age z-score (i)	Height-for-age z-score <-2 (ii)	Weight-for-height z-score (iii)	Weight-for-height z-score<-2 (iv)	Hemoglobine level (v)	Hemoglobine level<10g/dl (vi)
Program impact	0.048	-0.019*	-0.005	-0.003	0.128***	-0.027***
	(0.042)	(0.010)	(0.039)	(0.006)	(0.041)	(0.009)
Mean Control	-1,166	0.280	-0.022	0.078	11,217	0.169
Branch FE	YES	YES	YES	YES	YES	YES
Observations	10,570	10,570	10,175	10,175	10,568	10,568
R-squared	0.009	0.009	0.021	0.017	0.053	0.043

Notes: *Program impact* measures the coefficient on the assignment to treatment indicator, from a standard OLS regression. Branch fixed effects are included in every regression. There are 12 branches in the sample. Robust standard errors in parentheses, clustered at the cluster level. There are 214 clusters in the sample. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 5: Program Impact on CHP Interaction and Health Knowledge

Dependent variable	HH visited by a CHP in last 30 days (i)	Diarrhea from drinking untreated water (ii)	Zinc is effective against diarrhea (iii)	Mosquito bites are the only cause of malaria (iv)	Aware of food with added nutrients (v)	Bednets can help prevent malaria (vi)	Women should deliver at hospital (vii)	Average standardized effect (ii) - (vii) (viii)
Program impact	0.175***	0.041***	0.036***	0.027***	0.047***	0.001	0.000	0.064***
	(0.021)	(0.012)	(0.012)	(0.009)	(0.016)	(0.002)	(0.001)	(0.014)
Mean Control	0.054	0.373	0.227	0.071	0.591	0.991	0.997	
Branch FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7,018	7,018	7,018	7,018	7,018	6,977	7,018	
R-squared	0.158	0.035	0.084	0.056	0.065	0.005	0.005	

Notes: *Program impact* measures the coefficient on the assignment to treatment indicator. Dependent variables are indicators taking value one if: (i) household was visited by a CHP in the previous 30 days; (ii) respondent knows that diarrhea is transmitted by drinking untreated water; (iii) respondent believes that Zinc is effective in treating diarrhea; (iv) respondent believes that mosquito bites are the only cause of malaria; (v) respondent has ever heard of food with added vitamins or nutrients; (vi) respondent believes that bednets can help prevent catching malaria; (viii) respondent believes a woman giving birth should deliver at an hospital or health facility. Results in columns (i) to (vii) are obtained from a standard OLS regression. Column (viii) reports average (standardized) effect size across outcomes (ii) to (vii), using the seemingly-unrelated regression framework to account for covariance across estimates. Branch fixed effects are included in every regression. There are 12 branches in the sample. Robust standard errors in parentheses, clustered at the cluster level. There are 214 clusters in the sample. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 6: Program Impact on Health Behavior and Morbidity

Dependent variable	Treat water before drinking (i)	Child under treated bednet last night (ii)	Child ever received Vitamin A (iii)	Child had malaria over last 3 months (iv)	Child was treated with ACT for > 3 days (v)	Child had diarrhea over last 3 months (vi)	Child was treated with ORS/Zinc (vii)	Average standardized effect (i)-(vii) (viii)
Program impact	0.038**	0.051***	0.001	-0.013	0.004	0.005	0.053***	0.043***
	(0.015)	(0.014)	(0.012)	(0.014)	(0.015)	(0.009)	(0.020)	(0.013)
Mean Control	0.774	0.402	0.730	0.495	0.668	0.240	0.328	
Branch FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7,013	10,953	10,953	10,931	5,422	10,934	2,686	
R-squared	0.190	0.227	0.006	0.057	0.016	0.018	0.019	

Notes: *Program impact* measures the coefficient on the assignment to treatment indicator. Dependent variables are indicators taking value one if: (i) respondent treats the water before drinking it; (ii) the child slept under a treated bednet during the previous night; (iii) the child ever received a Vitamin A dose; (iv) the child ever fell sick with malaria during the previous 3 months; (v) the child that fell sick with malaria was treated with ACT drug for (at least) 3 days; (vi) the child ever fell sick with diarrhea during the previous 3 months; (vii) the child that fell sick with diarrhea was treated with ORS/Zinc. Results in columns (i) to (vii) are obtained from a standard OLS regression. Column (viii) reports average (standardized) effect size across outcomes (i) to (vii), using the seemingly-unrelated regression framework to account for covariance across estimates. Branch fixed effects are included in every regression. There are 12 branches in the sample. Robust standard errors in parentheses, clustered at the cluster level. There are 214 clusters in the sample. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 7: Program Impact on Health Visits

Dependent variable	Follow up visit...					
	...in first week after delivery (i)	...after child sick with malaria (ii)	...after infant sick with malaria (iii)	...after child sick with diarrhea (iv)	...after infant sick with diarrhea (v)	Average standardized effect (vi)
Program impact	0.081***	0.061***	0.073***	0.043**	0.081**	0.248***
	(0.020)	(0.014)	(0.028)	(0.017)	(0.037)	(0.066)
Mean Control	0.114	0.084	0.067	0.069	0.077	
Branch FE	YES	YES	YES	YES	YES	YES
Observations	1,925	5,335	631	2,228	408	
R-squared	0.074	0.096	0.147	0.077	0.144	

Notes: *Program impact* measures the coefficient on the assignment to treatment indicator. Dependent variables are indicators taking value one if the household received a follow up visit by an health care provider or community health worker: (i) in the first week after delivery; (ii) after a child under-5 fell sick with malaria; (iii) after a child under-1 fell sick with malaria; (iv) after a child under-5 fell sick with diarrhea; (v) after a child under-1 fell sick with diarrhea. Results in columns (i) to (v) are obtained from a standard OLS regression. Column (vi) reports average (standardized) effect size across outcomes (i) to (v), using the seemingly-unrelated regression framework to account for covariance across estimates. Branch fixed effects are included in every regression. There are 12 branches in the sample. Robust standard errors in parentheses, clustered at the cluster level. There are 214 clusters in the sample. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 8: Population Data and Flows

	Intervention group (115 clusters)	Control group (99 clusters)	p-value
Rate of in-migration	0.16 (0.12)	0.15 (0.11)	0.478
Rate of out-migration	0.07 (0.13)	0.07 (0.13)	0.991
Share of migrants	0.14 (0.09)	0.13 (0.08)	0.614

Notes: Data are mean (SD) estimated by combining data from baseline census, endline census, and endline sample household survey. P-values are adjusted for the stratified randomized design. Rate of in-migration is i_j/b_j and rate of out-migration is o_j/b_j, where $i_j = \hat{\theta}_j \times e_j$, $o_j = b_j - (i - \hat{\theta}_j) \times e_j$, b_j is number of households residing in cluster j at baseline, e_j is number of households residing in cluster j at endline, and the share of migrants $\hat{\theta}_j$ is an estimate of the share of households in cluster j that moved in to the cluster during the trial period, out of the total number of households living in the cluster at endline, based on the sample household survey.

APPENDIX

Table A.1: Under-5 mortality by wealth quartiles

	Intervention group (3,790 households)	Control group (3,228 households)
Quartile I		
Years of exposure to the risk of death	3,547	3,121
Reported deaths under-5	57	58
Mortality rate per 1000 years of exposure	16.1	18.6
Quartile II		
Years of exposure to the risk of death	2,918	2,750
Reported deaths under-5	42	53
Mortality rate per 1000 years of exposure	14.4	19.3
Quartile III		
Years of exposure to the risk of death	3,075	2,500
Reported deaths under-5	42	48
Mortality rate per 1000 years of exposure	13.7	19.2
Quartile IV		
Years of exposure to the risk of death	2,724	2,321
Reported deaths under-5	41	45
Mortality rate per 1000 years of exposure	15.1	19.4

Data are n and mortality rates from endline sample household survey. Wealth has been computed combining eight variables capturing ownership of durable assets (two sets of clothes for each household member, mobile phone, radio and television), infrastructure and housing characteristics (electricity, roof and floor material) and consumption habits (number of meals containing fish or meet served in a week), using Principal Component Analysis (PCA). The wealth index increases moving from quartile I to quartile IV. For 22 households asset information is missing.
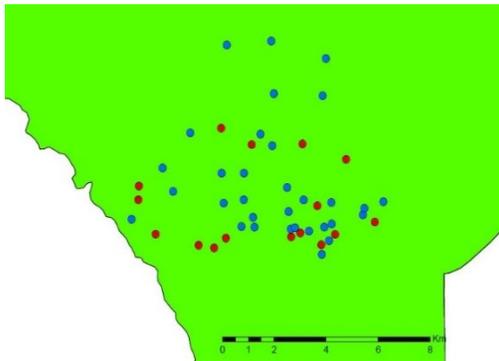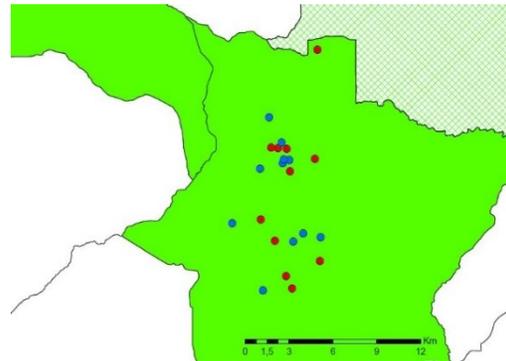
Figure A.1: Maps by Study District
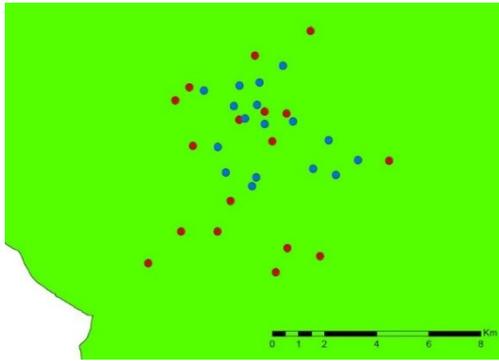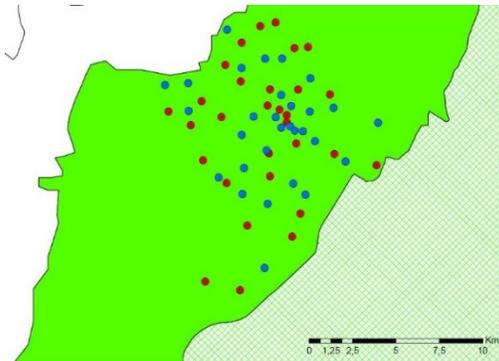

(a) Arua District
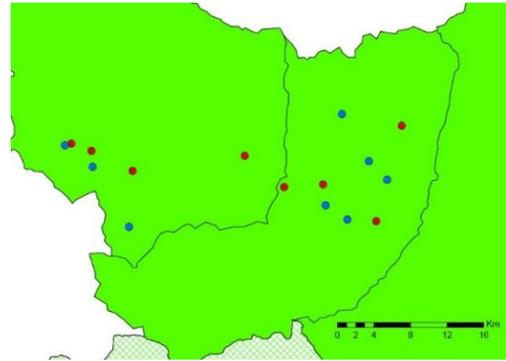
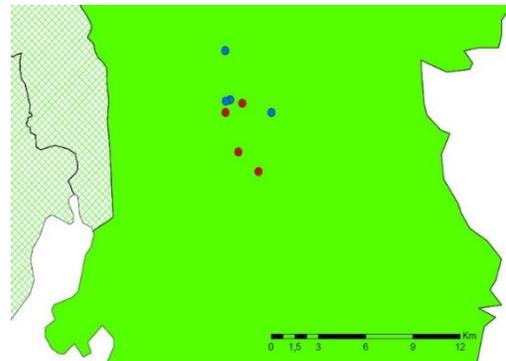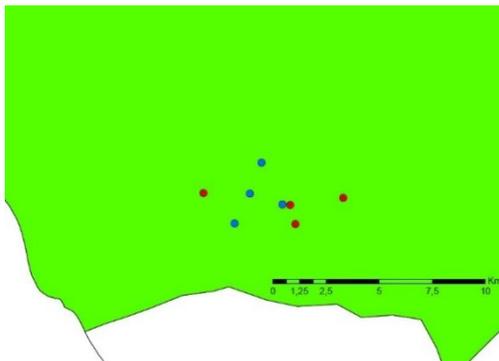**(b)** Bushenyi (West)/Sheema (East) Districts

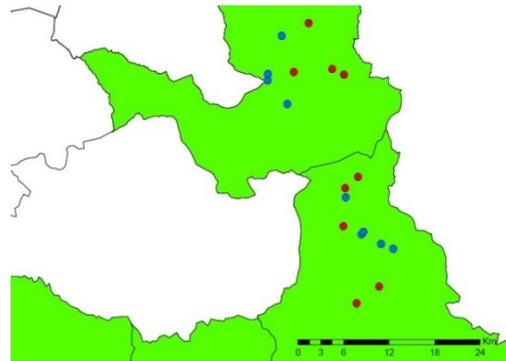**(c)** Jinja District

(d) Mbale District

**(e)** Mpigi District

(f) Mukono District

(g) Pallisa District

(h) Ibanda (North)/Mbarara (South) Districts

Notes: These figures are expansions of the map reported in figure 1. Green fully-colored indicate districts that were part of the study. Red and blue dots indicate respectively control and intervention villages included in the study.

Figure A.2: Computation of the months of exposure to the risk of death under-5 during the trial period
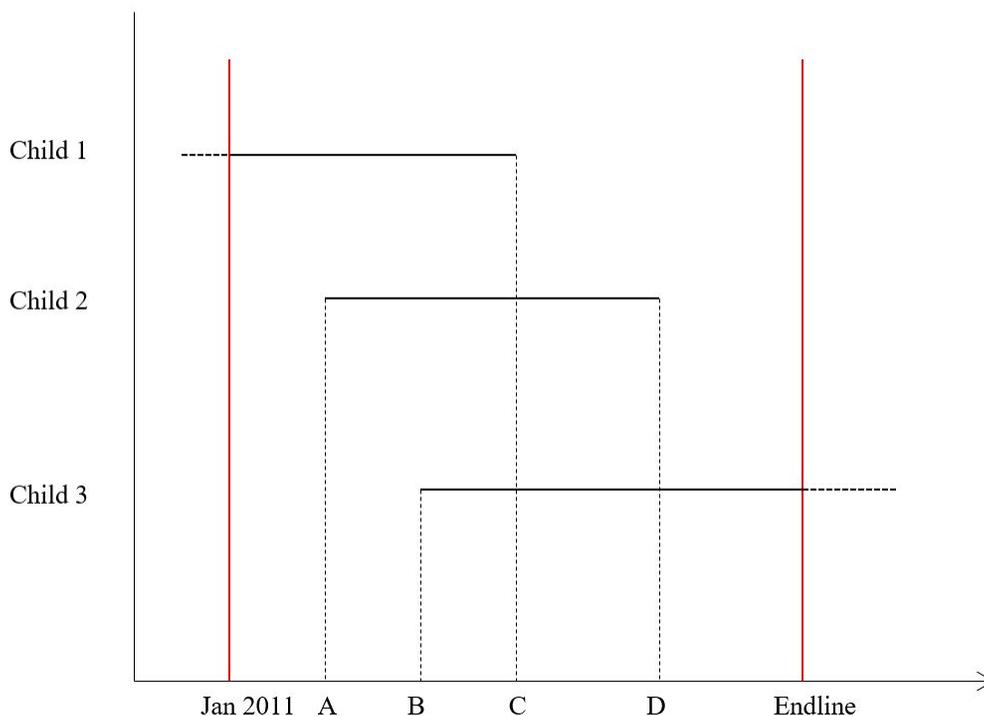
Notes: For each child, the number of month of exposure to the risk of death under 5 during the trial period is computed as the number of months between the birth date of the child, or the start date of the trial (January 2011) if the child was born before that date, and the date that the child turned five years if that occurred during the trial period, or the date of the endline household survey if the child was less than five years old at that time, or the date of the death of the child. The figure illustrates these different possibilities using the example of three children: child 1 was born before January 2011 and turned five at time C (the same computation would hold if the child died under age 5 at time C). Hence the exposure to the risk of death under 5 for child 1 is represented by the (rounded) number of months between January 2011 and time C. Child 2 was instead born at time A, during the trial period, but died at time D. Hence, in this case the exposure to the risk of death under 5 is represented by the (rounded) number of months between time A and D. Finally, child 3 was also born during the trial, at time B, and was still alive at the time of the endline. In this case the exposure to the risk of death under 5 is represented by the (rounded) number of months between time B and the time of the endline survey.