

Cash Transfers and Health

Evidence from Tanzania

David K. Evans

Brian Holtemeyer

Katrina Kosec



WORLD BANK GROUP

Africa Region

Office of the Chief Economist

November 2016

Abstract

How do conditional cash transfers impact health-related outcomes? This paper examines the 2010 randomized introduction of a program in Tanzania and finds nuanced impacts. An initial surge in clinic visits after 1.5 years—due to more visits by those already complying with program health conditions and by non-compliers—disappeared after 2.5 years, largely due to compliers reducing above-minimal visits. The study finds significant increases in take-up of health insurance and the likelihood of seeking treatment when ill.

Health improvements were concentrated among children ages 0–5 years rather than the elderly, and took time to materialize; the study finds no improvements after 1.5 years, but 0.76 fewer sick days per month after 2.5 years, suggesting the importance of looking beyond short-term impacts. Reductions in sick days were largest in villages with more baseline health workers per capita, consistent with improvements being sensitive to capacity constraints. These results are robust to adjustments for multiple hypothesis testing.

This paper is a product of the Office of the Chief Economist, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at devans2@worldbank.org and k.kosec@cgiar.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Cash Transfers and Health: Evidence from Tanzania

David K. Evans
The World Bank

Brian Holtemeyer
IFPRI

Katrina Kosec*
IFPRI

JEL Classification: I18, I13, O12, H42

Keywords: Cash transfers, health, government policy, health insurance

*This study benefitted at various stages from experts at the World Bank, the International Food Policy Research Institute (IFPRI), the Tanzania Social Action Fund (TASAF) and elsewhere. At TASAF, the evaluation has been supported by the Executive Director Ladislaus Mwamanga, as well as the former Executive Director Servacius Likwelile. Amadeus Kamagenge has led TASAF input to the evaluation, and his entire team has contributed with substantive and logistical support. We are also grateful to Harold Alderman, Domenico Fanizza, Margaret Grosh, Melissa Hidrobo, Jef Leroy, David McKenzie, Berk Özler, Anna Popova, Dena Ringold, Shalini Roy, and Magnus Saxegaard for comments and helpful discussions. We received financial support from the CGIAR Research Program on Policies, Institutions, and Markets, the International Initiative for Impact Evaluation (3ie), the Strategic Impact Evaluation Fund (SIEF), and the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD). Any comments or suggestions are welcome; the authors may be contacted at k.kosec@cgiar.org and devans2@worldbank.org.

1 Introduction

What role can cash transfers conditioned on health-seeking behavior play in alleviating the burden of poor health and limited access to formal medical care in Sub-Saharan Africa? Meta-analyses of programs from around the world suggest that conditional cash transfers (CCTs) can effectively alleviate extreme poverty and improve a range of human capital outcomes for children, at least during the period that the program is in place ([Fiszbein and Schady, 2009](#); [Leroy et al., 2009](#); [Independent Evaluation Group, 2011](#)). As the evidence base has grown, countries have raced to adopt CCTs. Almost every country in Latin America now has a CCT program ([Fiszbein and Schady, 2009](#)). Further, as of 2010, at least 14 countries in Sub-Saharan Africa had implemented a CCT program ([Garcia and Moore, 2012](#)). While there is considerable evidence about the impacts of cash transfers—conditional and unconditional—on education outcomes, less is known about their impacts on health. A 2014 global review found 142 studies showing the impact of cash transfers on education outcomes, but only 41 showing impacts on health and nutrition outcomes ([Andrews et al., 2014](#)).

Nowhere is the global burden of disease greater than in Sub-Saharan Africa. Life expectancy has increased by 20 years globally since 1970, but by only 10 years in Sub-Saharan Africa ([Institute for Health Metrics and Evaluation, 2013](#); [World Bank, 2013](#)). Further, Africa has a lower growth elasticity of poverty than any other region ([World Bank, 2013](#)). The region’s health problems are partly due to a pervasive lack of health investment by the public and private sectors, resulting in limited access to doctors and health facilities.

Existing literature on the health impacts of cash transfers yields mixed results. There are several program types: unconditional cash transfers (UCTs) and CCTs conditioned on health, education, or both. Especially in Sub-Saharan Africa, conditions are often “soft” (e.g., warnings or small penalties are applied instead of withholding the full transfer), making the difference between CCTs and UCTs less stark ([Garcia and Moore, 2012](#)). Relatedly, [Paxson and Schady \(2010\)](#) find that 28 percent of those participating in a UCT program in Ecuador thought conditions applied. However, even among cash transfer programs of the

same type, there are often mixed impacts on health. Among UCTs, [Haushofer and Shapiro \(2016\)](#) find that a transfer program in Kenya improved mental health and increased food consumption. UCTs have been found to improve anthropometric outcomes for girls—though not boys—in South Africa ([Duflo, 2003, 2000](#)). However, [Paxson and Schady \(2010\)](#) find null overall impacts of a UCT in Ecuador on cognitive, behavioral, and physical outcomes for children, with only those in the bottom expenditure quartile benefiting. And [Handa et al. \(2015\)](#) find no overall impacts of a UCT in Zambia on maternal health care utilization, with positive impacts only for women with better access to health services. CCTs conditioned on health have improved early childhood cognitive development ([Macours et al., 2012](#)) and child nutritional status ([Maluccio and Flores, 2005](#)) in Nicaragua, improved child nutritional status for a select group of children (younger children from rural areas) in Colombia ([Attanasio et al., 2005](#)), and had null impacts in Brazil ([Morris et al., 2004](#)) and Honduras ([Hoddinott, 2010](#)).¹ Finally, several studies explicitly compare the impacts of UCTs versus CCTs. [Akresh et al. \(2014\)](#) find an increase in preventative health care visits for a CCT in Burkina Faso conditioned on education and health, but not for a comparison UCT arm. [Robertson et al. \(2013\)](#), in contrast, do not find systematic benefits of CCTs over UCTs in Zimbabwe. And [Benhassine et al. \(2015\)](#) find that adding formal conditions to a labeled cash transfer (LCT) in Morocco—not subject to education conditions, but explicitly labeled as an education support program—may have *decreased* the overall impact of the program on school participation and learning. Similarly, for a CCT in Malawi conditioned on education, [Baird et al. \(2011\)](#) find that the UCT arm saw a greater reduction in teenage pregnancy among girls who had dropped out of school than did the CCT arm. And in the same study context, [Baird et al. \(2013\)](#) find greater mental health improvements and increases in usage of shoes among girls enrolled in school in the UCT arm than in the CCT arm. Given inconclusive findings in the literature on the health benefits of cash transfers, even among

¹A few studies examine more specialized CCTs. Interventions in India and Nepal offered incentives for maternal health investments, with mixed results ([Powell-Jackson et al., 2015](#); [Powell-Jackson and Hanson, 2012](#)). And interventions in Tanzania and Lesotho have provided incentives to remain free of sexually transmitted diseases, with positive outcomes ([Bjorkman Nyqvist et al., 2015](#); [De Walque et al., 2014](#)).

programs with similar designs, there is a need for greater understanding of the mechanisms through which cash transfers impact health.

We examine the impacts of a 2010 pilot CCT program in rural Tanzania on a range of health investments and outcomes. Among 80 study villages, 40 were randomly assigned to receive the CCT program, allowing us to estimate its causal impacts. Beneficiaries included both children aged 0–15 and elderly individuals aged 60 and older. Conditions of the program included visits to health clinics by young children aged 0–5 and by the elderly. Households were surveyed at baseline in 2009, again in 2011 after 18-21 months (about 1.5 years) of transfers, and finally in 2012 after 31-34 months (about 2.5 years) of transfers.

We find nuanced impacts of the CCT program. An initial surge in clinic visits after 1.5 years—due to more visits by both those already complying with program health conditions and non-compliers—disappeared after 2.5 years, largely due to compliers reducing above-minimal visits. We also find significant increases in take-up of health insurance. After 2.5 years, the program made households in treatment villages 36 percentage points more likely to participate in the government-run health insurance program (the Community Health Fund, or CHF) and raised the likelihood of financing treatment with health insurance by 16 percentage points. These impacts on health insurance are particularly interesting. Little previous work has examined the impact of cash transfers on participation in health insurance programs. Evidence from Mexico suggests that participation in a CCT program increased participants’ awareness that they were enrolled in a health insurance program, but in that case, actual enrollment was automatic upon enrollment in the cash transfer program ([Biosca and Brown, 2014](#)). The CCT additionally increased the likelihood of seeking treatment when ill. This latter result is important given recent research showing that timely clinic attendance when ill improves child health outcomes in Tanzania ([Adhvaryu and Nyshadham, 2015](#)). We also find that the program led to significantly higher investments in preventative health measures, including an 18 percentage points increase in shoe ownership, which the public health community associates with lower exposure to helminths ([Mascarini-Serra et al.,](#)

2011; Birn and Solórzano, 1999). Health improvements were concentrated among young children aged 0–5, with no detectable health improvements for elderly individuals similarly required to visit health clinics. Further, health improvements took time to materialize; we observe no improvements after 1.5 years, but 0.76 fewer sick days per month for 0–5 year olds after 2.5 years.² This suggests the importance of looking beyond very short-term impacts. Reductions in sick days were largest in villages with more baseline health workers per capita, consistent with improvements being sensitive to capacity constraints.

Overall, this evidence suggests a variety of mechanisms through which cash transfers may help to lift the burden of disease in Sub-Saharan Africa. We further show that these results are robust to adjustments for multiple hypothesis testing, estimation of linear as well as non-linear models, and both intent-to-treat and treatment on the treated estimates.

The remainder of the paper is organized as follows: Section 2 provides background information on health and the health care system in Tanzania, as well as the health conditions of Tanzania’s pilot CCT program. Section 3 describes the evaluation design, data, and outcomes of interest. Section 4 presents our empirical specification, the groups over which we examine heterogeneous treatment effects, balance tables showing the outcome of our randomization, and analysis of attrition. Section 5 characterizes our main empirical results and several robustness checks. Section 6 considers how our main impacts vary across different types of villages and households, and what this implies for the mechanisms likely driving treatment effects. Section 7 concludes.

2 Background

2.1 Health care and health in rural Tanzania

Tanzania is, in many respects, close to the Africa regional average in terms of health statistics. In 2012, 17.3 percent of the population contracted malaria versus 18.6 percent in Africa as a

²When we refer to illness in the last month, we are in all cases referring to the last four weeks.

whole. Likewise, 3.1 percent of the population was HIV positive, versus 2.8 percent in Africa. Life expectancy at birth is 61 years versus 58 for Africa. Yet on some measures, Tanzania diverges significantly from the rest of the region. Its under-five mortality rate (5.4 percent of live births) is just over half that of Africa as a whole (9.5 percent). Its maternal mortality ratio is almost 20 percent lower than that of Africa. Yet the health workforce is weaker in Tanzania, with just 0.1 doctors and 2.4 nurses and midwives per 10,000 population (versus an average of 2.6 and 12.0, respectively, for Africa) ([World Health Organization, 2014](#)).

Recent evidence from Tanzania demonstrates significant health improvements for children utilizing formal public health facilities ([Adhvaryu and Nyshadham, 2015](#)). In the early 1990s, the Tanzanian government introduced a health insurance program called the Community Health Fund (CHF). It is a voluntary prepayment scheme; members pay a fixed annual fee of 5,000 – 10,000 Tanzanian shillings (\$3 – \$6 US),³ depending on the region. Their entire family is then exempt from co-payments for visits to primary health care facilities ([Marriott, 2011](#)).⁴ As Tanzania’s CHF cross-subsidizes more costly to reach rural areas, it provides not only a risk-coping strategy, but also significantly reduces total out-of-pocket health expenses by the poorest (such as CCT program beneficiaries in our study regions) ([Ekman, 2004](#); [Mtei et al., 2007](#)). This is especially so since the poorest are often credit constrained. Nonetheless, 10 years after the introduction of the CHF, only 10 percent of Tanzanians were enrolled; one of the reasons cited was inability to pay ([Kamuzora and Gilson, 2007](#)).

2.2 Pilot CCT program

Tanzania’s pilot CCT program, implemented by the Tanzania Social Action Fund (TASAF, a social fund agency of the Tanzanian government), began delivering transfers in January of 2010. Its aims were to increase investments in health for young children (ages 0–5) and the elderly (ages 60 and over) and to increase educational investments for children aged 7–15. It

³In 2009 the exchange rate ranged from 1,280 to 1,467 per U.S. dollar ([Bank of Tanzania, 2015](#)).

⁴Up to 7 family members are exempt from co-payments—though tests and medications are subject to fees. Upon introduction of the CHF, child and maternal health services were already exempt from co-payments according to official government policy ([Babbel, 2012](#)).

operated in three districts—Bagamoyo (70 km from Dar es Salaam), Chamwino (500 km), and Kibaha (35 km)—where 80 eligible study villages were randomized into treatment and control groups of 40 villages each, stratified on village size and district (median village size at baseline was 560 households). Randomization was carried out after identification of potential beneficiary households in all 80 villages. At village meetings held prior to randomization, TASAF communicated that control villages would receive the program in late 2012, and the program would continue in treatment villages. Median village size was quite small (560 households at baseline, in 2009), and every village had both a primary school and a public dispensary or health center, facilitating fulfillment of program conditions.

Treatment households received transfers every two months. Transfer amounts ranged from US \$12 to US \$36, depending on household size and composition. The CCT provided US \$3 per month for orphans and vulnerable children up to age 15 (approximately 50 percent of the food poverty line) and US \$6 per month for vulnerable individuals age 60 or older. In our follow-up surveys, the median size of the last transfer is US \$14.12; assuming six annual payments of this size, this is about 13 percent of annual household expenditures.

While CCT payments were made at the household level, conditions applied at the individual level.⁵ Children aged 0–5 had to visit a health clinic at least six times per year (the condition was relaxed for children aged 2–5 to two visits per year starting in 2012),⁶ those age 60 or over had to visit at least once per year, and no health conditions applied to others. Both preventive and curative visits fulfilled the health clinic visit conditions of the program, though visits had to be to a public facility (either a dispensary, health center, or hospital). There were no further restrictions on the timing of visits, nor on the services to be received. Children aged 7–15 had to enroll in school and maintain an 80 percent attendance record.

⁵We lack administrative data on compliance with conditions. However, in each follow-up survey, we asked: “[For your last transfer payment,] did you receive less money than you usually get?” and “What do you think was the reason?” While one may hesitate to admit to non-compliance (for fear of sanction) and while this cannot tell us how many households had at least one payment reduced (it only tells us about the last), this gives some indication of compliance levels. At midline and endline respectively, 1.9 and 3.0 percent of treatment households reported receiving less than usual for a reason related to not meeting conditions.

⁶As our endline survey was carried out during August–October 2012, we define compliance with clinic visit conditions at endline for 2–5 year olds as having two or more clinic visits in the last year.

TASAF worked with an elected community management committee (CMC) in each village to select beneficiary households.⁷ The CMC surveyed the poorest half of households, collecting data on eight household characteristics: roof material, light supply, water supply, type of toilet, ownership of four different assets (vehicle/motorcycle, radio, iron, poultry), number of windows on the house, household size, and number of meals eaten per day. TASAF then carried out a proxy means test to propose a ranking of households by poverty level, for CMC and village leader approval. On average, 23 percent of households became beneficiaries.

3 Evaluation Design and Data

3.1 Evaluation design

We evaluate the impacts of the CCT program using three waves of data on beneficiaries and would-be beneficiaries (no data were collected from those not selected to be beneficiaries). Table 1 presents the chronology of the program and impact evaluation. A baseline survey was carried out during January–May 2009 and payments began in January 2010. A midline survey was conducted during July–September 2011 (18–21 months after transfers began) and an endline survey was conducted during August–October 2012 (31–34 months after transfers began). The baseline survey included 1,764 households (a subset of beneficiary households) comprised of 6,918 individuals. The quantitative data collection was supplemented by two rounds of qualitative data collection (following the midline and endline surveys) employing focus group discussions and in-depth interviews.

3.2 Data and outcomes

In each of the three survey rounds, we collected individual-level data on total health clinic visits in the last year,⁸ ownership of protective footwear (shoes and slippers) by children,

⁷CMC elections occurred at village meetings; 10–14 members were elected, with secret ballots. To run, a candidate had to have received financial training and successfully managed a past TASAF-supported project.

⁸We lack data on whether such visits were for preventive or curative purposes, and only know the total.

health (whether an individual was ill in the last month, and for how many days in the last month they were unable to perform their normal daily activities due to illness), reported ability to perform ordinary activities (doing vigorous activity, walking uphill, bending over or stooping, walking more than 1 km, walking more than 100 meters, or using a bath or toilet), anthropometrics (height, weight, middle upper-arm circumference, and z-scores for height-for-age, weight-for-age, weight-for-height, and body mass index-for-age), and—for those ill in the last month—the location where medical care was sought (no treatment sought, public dispensary, public hospital, public health center, private pharmacy, traditional healer, private dispensary/ hospital/ clinic, or mission dispensary/ hospital).⁹ Among those seeking treatment, we further gathered data on health care financing methods (free treatment, loans, cash or assets, or health insurance). We also collected household-level data on expenditure on formal insurance¹⁰ and—at endline—whether the household participates in the CHF.

4 Methods and Empirical Strategy

4.1 Empirical specification

We carried out follow-up surveys in 2011 and in 2012 to capture both short-term (1.5 years) and medium-term (2.5 years) impacts of the program. Given random assignment to treatment, we recover causal intent-to-treat estimates from the following empirical specification:

$$h_{it} = \beta_0 + \beta_1 2011_t + \beta_2 2012_t + \delta_1 T_i \times 2011_t + \delta_2 T_i \times 2012_t + \alpha_i + \epsilon_{it} \quad (1)$$

where i indexes individuals and t indexes the survey round. h_{it} is a health-related outcome, α_i are individual fixed effects, $T_i=1$ in a village assigned to treatment and zero otherwise, $2011_t=1$ at the time of the midline survey (July–September 2011) and zero otherwise, and

⁹Data on protective footwear were only collected for children aged 0–18, data on anthropometrics for children aged 0–5, and data on ordinary activities for those aged 60 and over. Individuals sick in the last month were asked to report the primary health provider and payment method for their main health problem.

¹⁰Insurance expenditure data are unfortunately not further disaggregated by type of insurance.

2012_{*i*}=1 at the time of the endline (August–October 2012) and zero otherwise. When we consider a household-level outcome, i indexes households.

In treatment villages, 9.0% of households did not receive treatment—likely due to last-minute changes in community prioritization or household refusal. In control villages, 0.6% of households received treatment—likely due to their proximity to a treatment village. As a result, our intent-to-treat estimates represent a lower bound on the actual impact of receiving transfers. We also estimate the effect of treatment on the treated by using the fitted values from a regression of treatment on assignment to treatment in place of T_i in equation 1.¹¹

4.2 Heterogeneous treatment effects examined

We estimate the overall impacts of the CCT program as well as its impacts on several sub-groups. First, we examine impacts by age group. As health conditions applied only to children aged 0–5 and elderly aged 60 and over, and given that each of these two age groups has a different set of health issues and faced different conditions under the CCT program, it is instructive to examine program impacts on them separately. Overall impacts include all individuals in the surveyed households, not only all individuals in the two sub-groups.

Second, for two central outcomes likely to be heavily influenced by the quality of available health care—health clinic visits and health during the last month—we examine heterogeneous impacts of the CCT program by baseline health clinic staff (doctors, nurses, and other assistants) per capita. Specifically, we divide villages into two types: those with above-median and below-median health clinic staff per capita at baseline. This helps us assess if improvements are sensitive to capacity constraints

Finally, for outcomes likely to be influenced by how credit constrained a household is—shoes and slipper ownership, expenditure on insurance, participation in the CHF, whether one treats illness and where (public or private facilities), and how one finances treatment—we examine heterogeneous impacts by baseline household asset wealth. This allows us to

¹¹We use the Stata package `xtivreg` written by [Schaffer \(2010\)](#).

observe how the program affects the moderately poor (top half of beneficiaries in terms of asset wealth) versus the extremely poor (bottom half). To capture asset wealth, we carry out a principal components analysis (PCA) using dummy variables for ownership of 13 assets.¹²

4.3 Outcome of the randomization

A comparison of baseline sample means in treatment and control villages reveals balance on most outcomes (Appendix Table A1, Panel A). Across 41 outcomes, for only six (three) are there significant differences at the 10 percent (5 percent) level. There are no overall differences in health between treatment and control villages, but 0–5 year olds in treatment villages were 6 percentage points more likely to be ill or injured and had 0.51 more sick days in the last month than those in control villages. Weight-for-age z-scores were slightly lower in treatment villages. And 0–18 year olds in treatment villages were 10 percentage points less likely to own shoes than were those in control villages. Households in treatment villages also spent slightly more on insurance. However, we see balance on all other outcomes of interest. Table A1, Panel B shows similar balance on individual and household demographic characteristics and village characteristics. The only significant difference is that treatment households are less likely to have an improved floor. We use individual (or household, for outcomes that vary at that level) fixed effects to account for baseline imbalances.¹³

4.4 Attrition

If attrition were correlated with treatment status, one might worry that attrition had compromised the internal validity of the results.¹⁴ Fortunately, this is not the case, as shown in Table A2. Columns (1)–(4) consider household attrition, columns (5)–(8) consider individual

¹²These include whether the household owns an iron, refrigerator, television, mattress or bed, radio, watch or clock, sewing machine, stove, bicycle, motorcycle, car or truck, wheelbarrow or cart, and mobile phone.

¹³In the case of anthropometric outcomes, we use village \times cohort fixed effects. These results are robust to instead using individual fixed effects.

¹⁴Between baseline and midline, 8.6 percent of households attrited from the sample, and between baseline and endline, 13.2 percent of households attrited.

attrition, and columns (9)–(12) consider individual attrition for those for whom health conditions applied (children aged 0–5 and those aged 60 and over). For each of the three analyses, we consider attrition at midline and at endline. Odd-numbered columns regress a dummy for attrition on our treatment dummy, while even-numbered columns regress a dummy for attrition on our treatment dummy, an array of controls (gender, age, age-squared, a dummy for having some education, and a household asset index), and the interactions of these controls with treatment. Where we examine household attrition, we use the values of these controls for the household head; where we examine individual attrition, we use the values for the head as well as the individual. In no case does the treatment dummy significantly predict attrition. F-statistics for the joint significance of the treatment dummy and the interaction terms further indicate that these coefficients are never jointly significant. Overall, we conclude that attrition does not affect the internal validity of our results.

5 Results

5.1 Health clinic visits

At baseline, the average 0–5 year old visited a clinic 8.3 times per year (compared to the program condition of six visits), and the average individual aged 60 or older visited 2.8 times (compared to the program condition of one visit). This reveals that on average, individuals were exceeding program conditions at baseline.¹⁵ Universal compliance with the program could thus occur at follow-up even with a zero net increase in clinic visits if more frequent visits by those not previously in compliance were offset by less frequent visits by those already complying at baseline. In Table 2, we examine the impact of the CCT program on the number of health clinic visits in the last year. We focus on overall impacts, impacts on children aged 0–5, and impacts on those age 60 and over. (At endline, we have clinic visit

¹⁵At baseline, 59 percent of children aged 0–5 and 65 percent of those aged 60 and over met or exceeded required health clinic visits, making program conditions non-binding for them.

data for all individuals, while at midline, we only have it for these two groups.)

At midline (1.5 years after treatment began), treatment led to 2.3 more visits (preventive or curative) per year for children aged 0–5 (column 2) and 1.1 more visits for those aged 60 and over (column 3). Relative to the baseline mean number of visits for each age group, these represent increases of 28 and 39 percent, respectively, which are comparable to findings in the literature. For example, [Levy and Ohls \(2007\)](#) find that a CCT program in Jamaica increased preventive health center visits by children aged 0–6 by 0.28 visits every six months (a 38 percent increase relative to the baseline mean), while [Akresh et al. \(2014\)](#) find that a CCT program in Burkina Faso increased annual preventive care visits by children aged 0–5 by 0.43 visits (a 49 percent increase relative to the control group mean). These statistically significant effects, however, disappear at endline (2.5 years after treatment began) for both age groups.¹⁶ The results are robust to instead estimating a Poisson model that accounts for health clinic visits being a count data outcome (Appendix Table A9)¹⁷ and to instead estimating the impact of treatment on the treated (Appendix Table A3).

In Table 3, we examine the impact of the CCT program on the rate of compliance with annual clinic visits conditions. While we lack administrative data on clinic visits, comparing self-reported visits over the last 12 months with program conditions for total annual visits is instructive. The conditions required six visits for those under age 2 and one visit for those over age 60; 2–5 year olds needed six visits at midline but two visits at endline. We see that treatment increased compliance with program conditions among both 0–5 year olds and those over aged 60 at midline. By endline, however, treatment increase compliance with program conditions only among those over aged 60, and not among children aged 0–5.

Several caveats and observations are warranted. First, households in control villages that by endline anticipated receiving the program within a few months¹⁸ may have increased

¹⁶At endline, those in treatment villages were still receiving the CCT program and expected it to continue. Those in control villages anticipated being enrolled within the next few months (by late 2012).

¹⁷Results hold whether we use heteroskedasticity robust or bootstrapped standard errors.

¹⁸The endline survey was carried out during August–October 2012, and control villages were told at baseline to anticipate receiving the program in late 2012.

clinic visits preemptively for fear of being cut from the list of targeted households. Indeed, when we consider compliance with clinic visit as our outcome (Table 3), we find a positive and highly statistically significant coefficient on the endline dummy for 0–5 year olds—consistent with an overall increase in childrens’ compliance with program conditions across both treatment and control villages at endline. Second, it is important to interpret these findings in light of high baseline rates of compliance that made the health visit conditions non-binding for many.¹⁹ Program emphasis on clinic visits may have increased the salience of health services and led households to initially increase visits despite the average household already satisfying visit conditions. Subsequently—by endline—individuals’ understanding of the conditions may have improved, and they may have reduced visits to only those that were necessary, still exceeding the program conditions on average. Third, health improvements due to the program that were realized by endline but not at midline—detailed in Section 5.4—may have reduced demand for clinic visits by endline. Finally, while we lack data on clinic service quality, it may have improved by endline, requiring fewer visits to receive similar care (e.g., receiving more and better services at a first visit could preclude the need for a follow-up visit). We present further evidence and discussion of why clinic visits may have increased at midline, but were subsequently unaffected at endline, in Section 6.

5.2 Protective footwear

While health clinic visits are an important aspect of individual investment in health, investments that individuals make to prevent health problems from occurring are also important. We examine the impacts of treatment on ownership of two types of protective footwear: shoes and "slippers" (i.e., open-toed footwear). Table 4 shows that the CCT program led to a significant, 18 percentage point increase in shoe ownership among 0–18 year old children by midline that persisted at endline (column 1).²⁰ A null impact on slipper ownership at midline

¹⁹While non-binding conditions make a CCT similar to a UCT or LCT, Section 1 discusses how UCTs are prevalent, and many CCTs have soft conditions. Thus, our study context is not atypical.

²⁰Estimates of the effect of treatment on the treated are similar (Appendix Table A4).

changed to a significant, 8 percentage point increase by endline (column 2). This suggests that the program did not lead to a substitution between shoes and slippers, but rather increased take-up of both products by endline. Further, impacts were largest for ownership of shoes—which provide better protection. These impacts are remarkable considering baseline ownership rates of shoes and slippers were only 42 percent and 63 percent, respectively.

5.3 Health insurance

We also examined program impacts on take-up of health insurance. As we discuss in Section 2, participation in Tanzania’s government-run health insurance program, the CHF, should not only help households cope with the risk of health shocks, but also reduce out-of-pocket expenditures on health given cross-subsidies built into the scheme, favoring the rural poor. Table 4 shows that treatment increased household expenditures on insurance sixfold by midline and eightfold by endline (column 3). It also increased participation in the CHF; while we lack baseline data on participation rates, by endline, households in treatment villages were 36 percentage points more likely to participate than were households in control villages (column 4). This is strikingly large given that at baseline, only 3 percent of individuals who sought treatment for illness in the last month reported using health insurance to fund it.

Table 5 examines the health care financing methods of individuals who reported being ill in the last month and treated the illness.²¹ We find that at midline, the program reduced payment for health care using cash or an asset by 18 percentage points, which is a 27 percent decrease from the baseline mean of 65 percent. This same effect size was sustained at endline. Those no longer financing treatment with cash and assets began using health insurance; we see a 16 percentage point increase in use of health insurance at midline to finance treatment, which swelled to 28 percentage points by endline (from a baseline mean of only 2.7 percent).

²¹In Appendix Table A5, we examine whether treatment affects selection into who reports being ill or injured in the last month. We find that few interactions of individual and household characteristics with treatment are statistically significant predictors of illness or injury. Further, when we conduct a test of the joint significance of these interaction terms for individuals for whom health conditions applied (those aged 0–5 and aged 60 and over), we find insignificance at both midline ($p = 0.369$) and endline ($p = 0.612$).

Our findings on health insurance may be driven by several factors. First, if liquidity has been a binding constraint, a CCT program may increase take-up of insurance.²² Second, the CCT program may have improved access to information about the CHF and lowered barriers to enrollment. In qualitative data, health clinic staff in treatment villages reported going to the place where beneficiaries collected transfers to tell them about the CHF and encourage sign-up while they still felt rich (Evans et al., 2014). Finally, other research on the same CCT finds that it increased familiarity with and trust in local leaders and health care providers (Evans et al., 2016). Combined with evidence that trust increases take-up of insurance (Dercon et al., 2015), this too may explain increased participation in the CHF.

5.4 Health

Table 6 reports the effects of treatment on two key health outcomes: whether or not an individual was ill or injured in the last month, and the number of days that the individual was unable to perform their normal daily activities in the last month due to illness (sick days) (Panel A). These capture, respectively, the extensive and intensive margins of illness. We see that at midline, treatment had no significant impact on either health outcome. However, at endline, treatment significantly reduced both the extensive and intensive margins of illness. In particular, for the sample as a whole, treatment resulted in a 4.3 percentage point reduction in the incidence of illness or injury in the last month (p-value = 0.101); while of borderline statistical significance, this is a sizeable 17 percent decrease relative to the baseline mean incidence of 27.6 percent. When we instead compute the effect of treatment on the treated (Appendix Table A6), we observe a statistically significant (p < 0.10), 4.6 percentage point reduction in incidence of illness or injury in the last month. For the sample as a whole, treatment also resulted in a statistically significant, nearly half-day decrease in sick days in the last month (a 27 percent decrease relative to the baseline mean of 1.64 sick

²²A desire to insure against health shocks can be understood in light of the frequency of such shocks in our study context; at baseline, 55 percent of households reported experiencing a health shock in the last five years (specifically, a chronic or severe illness or accident of a household member, or a death in the family).

days). These treatment effects seem to be strongly driven by health improvements for young children (ages 0–5), for whom the reduction in incidence of illness in the last month is 10.7 percentage points (significant at the 10 percent level) and the reduction in sick days is 0.76 (significant at the 5 percent level). We find no significant overall program impacts for those aged 60 and over, either on the extensive or the intensive margins.²³ Similar results hold when we instead estimate a Poisson model (Appendix Table A9). While the program has health benefits, these take time to materialize, are most prominently on the intensive rather than extensive margin of illness, and accrue predominately to young children.

Despite overall health improvements, the CCT program did not change the ordinary activities that elderly individuals could perform, as shown in Appendix Table A10. Specifically, it did not have significant impacts on individuals’ reported ability to do vigorous activities, walk uphill, bend over or stoop, walk more than 1 km, or use a bath or toilet, nor did it affect a simple 0–6 index of these activities (the “ordinary activities index”). One exception is the ability to walk more than 100 meters (a dummy that had a very high baseline mean of 0.96); there, we find a very small negative impact of the program at endline. Overall, however, the program did not have systematic impacts on the types of activities that individuals could perform; rather, it changed the number of days that they could perform their activities.

There are several reasons that health may have improved. First, given that the CCT program increased health clinic visits at midline, this increased health-seeking behavior may have itself improved health by endline. Second, additional income, insurance, and the added familiarity with health clinics generated by the program may have spurred individuals to visit clinics promptly whenever ill, thus reducing the duration of illness. We explore this possibility in Section 5.6; if clinic visits were better timed (even if their aggregate numbers did not increase), this might explain why our results on the intensive margin of illness are the most robust. Third, the program may have generally stimulated health-promoting investments by households. Existing research on the CCT program shows that it did not

²³Treatment did not affect rates of mortality or the number of household members in different age groups (Appendix Tables A7, A8, and Panel C of Appendix Table A9).

increase food consumption during the last week at either midline or endline, but that it did increase expenditures on non-food items in the last 12 months—including on women’s and children’s clothing—and increased the number of goats and chickens households owned (Evans et al., 2014). Further, the program increased children’s shoe ownership (Section 5.2).

The findings on improved health demonstrate the importance of taking care to evaluate health outcomes after an appropriate period of time, as advocated in general by King and Behrman (2009). At least in this study, positive health impacts do not appear after 1.5 years of transfers, though do appear after 2.5 years. In Section 6, we present further evidence and discussion on why health may have improved by endline but not midline, and what role program health conditions may have played in delivering health benefits.

5.5 Anthropometrics

Appendix Table A11 reports the effects of treatment on a number of anthropometric outcomes for children aged 0–5: height-for-age, weight-for-age, weight-for-height, body mass index (BMI)-for-age, height, weight, and middle upper-arm circumference (MUAC) (columns 1–7, respectively). These regressions use village \times 6-month age cohort fixed effects since very few children were in the 0–5 age range for multiple observations during 2009–2012. We find no evidence that treatment influences these outcomes.²⁴ The lack of anthropometric effects is striking; it contributes to a mixed literature on the impacts of CCTs on child anthropometrics (Fiszbein and Schady, 2009). The result is less surprising when considering the null impacts of the program on food consumption (Evans et al., 2014).

5.6 Health care provider type

Table 7 examines the impacts of the CCT program on health care provider decisions of individuals who reported being ill in the last month.²⁵ Individuals either fail to treat their

²⁴We also find null results when only considering children under age 2.

²⁵At baseline, 28 percent of individuals reported being ill or injured in the last month; 85 percent sought treatment for their illness—49 percent in the public sector and 36 percent in the private sector.

main health problem (15 percent did so at baseline) or visit one of several different types of public and private providers. As the health conditions of the program required visits to be at public facilities, we anticipated finding larger impacts of treatment on public than on private facility visits. We find that at midline, the program reduced failure to seek treatment by 12 percentage points (column 1). That illness was more likely to be treated may explain why health impacts were most robust on the intensive margin, with the CCT program generally contributing to shorter spells of illness. At both midline and endline, the program increased use of public dispensaries; by 17 percentage points at midline, and 15 percentage points at endline (column 2). However, it did not impact use of private providers (columns 5–8). This is consistent with the health clinic visit conditions of the program, which counted only visits to public facilities and not private facilities. Rather than drawing individuals from the private to the public sector, we see a shift from a failure to treat illness to treatment at a public dispensary. Further, the program did not lead individuals to treat illness at a public hospital or a health center (columns 3–4). These are larger facilities that would typically offer more services (and potentially more qualified staff), but which are usually further away.

5.7 Robustness: Corrections for multiple hypothesis testing

A growing literature recognizes the risk of finding false positives when testing multiple hypotheses and advances correction methods. Two popular methods are the Benjamini and Hochberg (BH) and the Benjamini–Krieger–Yekutieli (BKY) methods, which control for the false discovery rate (FDR) (Benjamini and Hochberg, 1995; Benjamini et al., 2006). We compute the q-values (i.e., p-values corrected for multiple testing) of each. As a third test, we apply a Bonferroni correction—a method of controlling the family-wise error rate (FWER) that involves multiplying each p-value by the number of tests performed. While simple to compute, it suffers from poor power (Anderson, 2008) and is often used as an upper bound on the FWER (Hochberg, 1988). We thus rely primarily on the BH and BKY results, but take the Bonferroni as a useful guide to the lower bound of the significance of our results.

Appendix Table A12 reports the resulting q-values from the three correction methods for all originally statistically significant impacts. For each method, a group of hypotheses is defined by the follow-up survey round (midline or endline) and broad type of outcome being considered (e.g., child anthropometrics or ordinary activities). This is usually equivalent to grouping together all of the hypotheses within a table for a given survey round.²⁶ Hypotheses associated with heterogeneous treatment effects are grouped with the hypotheses of overall treatment effects, despite being displayed in separate tables. In total, we observe 61 statistically significant impacts of the CCT program in the paper’s main tables. When we correct for multiple testing using the BKY (BH) method, 45 (43) remain significant. With the more conservative Bonferroni method, 35 remain significant.

Our main conclusions still hold. Under all three correction methods, the program significantly increased health clinic visits at midline but not endline. Furthermore, it boosted children’s shoe ownership and household expenditure on insurance at both midline and endline. Reductions in the intensive margin of illness by endline are still significant with both BKY and BH, though reductions in the extensive margin of illness do not survive these corrections. This may indicate that the CCT program’s ability to reduce illness is principally concentrated in its reduction of the severity of illness—and how debilitating it becomes—rather its incidence. Notably, under all three corrections, our finding that the program increases participation in the CHF remains statistically significant at the 1 percent level. Similarly, findings that the CCT program reduced the likelihood of failing to treat illness at midline, and increased visits to public dispensaries (but not private ones) at both midline and endline remain robustly significant under all three corrections. Finally, at both midline and endline, the finding that the program reduced use of cash or assets to finance health care and boosted usage of health insurance are again significant under all three corrections.

²⁶There are three exceptions. In Table 4, protective footwear outcomes are grouped separately from insurance-related outcomes—just as they are separately considered in Section 5. In Table 6, hypotheses related to the extension and intensive margins of health are grouped separately. And in Appendix Table A10, we omit the ordinary activities index from our grouping; estimation of impacts of treatment on an index simply serves as an additional check on the robustness of these null findings.

6 Mechanisms

We gain additional insight into the mechanisms likely driving the impacts of treatment by separately examining sub-groups of beneficiaries. We already considered impacts by age group. However, as we discuss in Section 4.2, we also consider two additional types of heterogeneous treatment effects. For two central outcomes likely to be heavily influenced by the quality of available health care—health clinic visits and health during the last month—we examine heterogeneous impacts on villages with above-median versus below-median health clinic staff per capita at baseline. This helps us assess if improvements are sensitive to capacity constraints. For outcomes likely to be influenced by how credit constrained a household is—shoes and slipper ownership, expenditure on insurance, participation in the CHF, whether one treats illness and where (public or private facilities), and how one finances treatment—we examine heterogeneous impacts in moderately poor households (top half of beneficiaries in terms of asset wealth) versus extremely poor households (bottom half).

6.1 Impacts by health clinic staffing levels

When we examine heterogeneous treatment effects by baseline health clinic staff per capita (Table 8), several interesting findings emerge. First, Panel A reveals that we cannot reject the null hypothesis that the CCT program had the same effect on health clinic visits in villages with few baseline health staff per capita (the bottom half of the distribution) as in villages with many (the top half). This is true overall and for both age groups (0–5 year olds and those age 60+). This provides suggestive evidence that the impacts of treatment on clinic visits would not be enhanced by increasing clinic staff per capita. Second, Panel B shows heterogeneous impacts on health by baseline clinic staff per capita. Here, we find that reductions in sick days are concentrated in villages with more health staff per capita, with no significant impacts in villages with few staff per capita. For individuals in villages that were highly-staffed at baseline, the average reduction in sick days in the last month is

0.96 (compared to an insignificant 0.07 days in more poorly-staffed villages). The difference between these two effects is statistically significant at the 5 percent level. This suggests that reductions in the intensive margin of illness may in fact be conditional on a village having sufficient staff to attend patients and treat illness. It is important to note, however, there are no differential impacts of the program by baseline staffing levels on sick days for children aged 0–5. Clinic staffing may matter more for older individuals—possibly as they are less integrated into the health system. Additionally, there are no differential impacts of treatment on the incidence (extensive margin) of illness by clinic staffing levels.

Appendix Figure A1 helps illustrate these impacts. Sub-figure (a) presents two village-level scatterplots—one for midline, one for endline—each with 80 data points: 40 show treatment village averages (in black circles), and 40 show control village averages (in gray squares). We plot on the y-axis the average change in the number of sick days in the last month between baseline and follow-up (positive numbers indicate an increase, negative a decrease) and on the x-axis village health clinic staff per capita at baseline. We include separate linear fits for treatment and control. At midline, we see that there are essentially no health improvements in either treatment or control villages, across the full distribution of clinic staff per capita. By endline, however, treatment villages on average have greater reductions in sick days than do control villages across the full range of values of staff per capita. However, the difference is greatest in villages with especially high staff per capita.

Sub-figure (b) of Appendix Figure A1 presents a similar analysis where the outcome is now the average change in the number of clinic visits in the last year between baseline and follow-up. At midline, treatment villages on average experience larger increases in clinic visits than do control villages across the full range of values of clinic staff per capita. While the difference is greater in villages with higher staff per capita, this difference was not statistically significant in regressions. At endline, we see similar reductions in clinic visits in treatment and control villages across the full distribution of staff per capita—possibly reflecting better average health at endline (as seen in sub-figure (a)) that alleviated capacity constraints.

6.2 Impacts by household wealth

Examining heterogeneous treatment effects by baseline household asset wealth (Table 9) reveals several interesting results. First, as shown in Panel A, we do not find significant differences in the impacts of the program on whether one treats illness (column 1), and whether or not they treat it in a public dispensary (column 2), according to baseline household wealth.²⁷ However, Panel B shows that the impacts of the CCT program on shoe ownership (column 1), slipper ownership (column 2), and insurance expenditures (column 3) are responsive to baseline household wealth. For each of the three, the impacts of treatment are larger for the extremely poor (those in the bottom half of asset wealth at baseline) than for the moderately poor for both follow-up survey rounds. These differences are in several cases statistically significant. At midline, the extremely poor saw a significantly greater increase in shoe ownership and insurance expenditure than did the moderately poor, while at endline, the extremely poor had a significantly greater increase in slipper ownership than did the moderately poor. The effect on CHF participation at endline is slightly larger (but not statistically significantly different) for extremely poor households (column 4). However, Panel C reveals that while the increase in use of health insurance at endline was larger for the extremely poor than for the moderately poor, the difference is not statistically significant. Overall, these results suggest that not only can a CCT program increase take-up of products that tend to prevent health problems from occurring and help households cope with health-related risks, but also that in some cases, the poorest of the poor benefit most.

6.3 Exploratory analysis: Timing and drivers of health impacts

Our findings raise two important and related questions that have not been fully answered: First, why do health improvements show up at endline but not at midline? Second, did clinic

²⁷While we find no significant impacts of treatment on other types of providers (columns 3–8), there are two statistically significant differences worth noting. First, at midline, treatment led the moderately poor to increase use of public health centers—though had no significant impact on their use by the extremely poor. Second, at midline, treatment led the extremely poor to decrease use of private dispensaries, hospitals, clinics, and stores—though had no significant impact on their use by the moderately poor.

conditions contribute to the health improvements realized from the program, or were they unnecessary? We carry out exploratory analyses to shed further light on both questions.

Appendix Figure A2 presents four village-level scatterplots, each containing 80 data points: 40 showing treatment village averages (in black circles), and 40 showing control village averages (in grey squares), with separate linear fits for each set of 40. We plot on the y-axis the average change in the number of sick days in the last month between baseline and follow-up (positive numbers indicate an increase in sick days, negative numbers a decrease). On the x-axis, sub-figure (a) has the village average number of clinic visits per person at baseline, while sub-figure (b) has the share of the village already complying with the health conditions at baseline. Each sub-figure features a plot for each of the two follow-up surveys.

Sub-figures (a) and (b) tell a similar story; reductions in the intensive margin of illness (sick days) show up at midline for a subset of the population (even if there are no significant reductions in the aggregate, as shown in our regressions): those already complying with or exceeding program health conditions at baseline and those with more clinic visits at baseline (call these “the compliers”). This suggests that it was not the conditions driving initial (midline) health improvements; it was the income effect of the transfer on individuals already visiting clinics frequently. However, by endline, the story reverses. It was those who were not complying with the conditions at baseline and those with fewer clinic visits at baseline (call these “the non-compliers”) who saw the greatest reductions in sick days.

Our regressions showed that treatment boosted health clinic visits at midline. Appendix Figure A3 further reveals that this midline increase in clinic visits was larger among non-compliers than among compliers.²⁸ This suggests that the CCT program helped individuals in two waves. First, compliers experienced a mild reduction in sick days at midline, likely due to income effects of the CCT. Second, non-compliers experienced health benefits, but with a lag. By midline they were induced to increase their health clinic visits—likely due to both the conditions of the program and the income from the transfer permitting them

²⁸This is also the case at endline, where despite no aggregate increase in clinic visits at endline due to treatment, we see increases for those with low rates of compliance with health conditions at baseline.

to go to the clinic when sick rather than allowing illness to go untreated. But it was not until endline that this paid dividends in terms of better health. Indeed, by endline, these non-compliers likely benefited from at least two factors: their increased exposure to clinics at midline as well as the fact that treatment had already reduced illness in a number of children in the village (i.e., the compliers), spurring an overall reduction in infectious disease rates that shows up in our regressions as statistically significant overall impacts of treatment on sick days at endline. Appendix Figure A2 also reveals that we see larger reductions in sick days at endline among those with fewer clinic visits at baseline (sub-figure (a)) and among those with lower rates of compliance with conditions at baseline (sub-figure (b)), consistent with conditions helping to explain overall health improvements.

7 Conclusion

This paper provides evidence that, after 2.5 years, a conditional cash transfer (CCT) program in Tanzania made children aged 0–5 experience fewer monthly sick days. We find no evidence of health improvements for those aged 60 and over despite their having also been required to visit a health clinic as a condition of the transfer, suggesting greater promise of such programs for the young.

The statistically significant improvements in health outcomes after 2.5 years are particularly striking given that the program’s initial effect of increasing annual clinic visits had disappeared after 2.5 years of transfers. If health improvements were not driven by increased clinic visits, then what was the cause? Previous analysis suggests that these study households did not significantly increase consumption (Evans et al., 2014). Instead, we show that households used their transfers to reduce the risk of high health care costs. Households invested in footwear for their children, which reduces exposure to health risks. Households were substantially more likely to invest in a government-run health insurance program. They went on to utilize that health insurance to finance clinic visits when ill. Although the total

number of clinic visits was not higher among beneficiary households, participation in the insurance program meant that those households could attend the clinic when they most needed it, rather than letting immediate financial liquidity determine when, in the course of an illness, to visit the clinic. This is consistent with findings from [Adhvaryu and Nyshadham \(2015\)](#), who demonstrate that households that access formal sector malaria treatment in a more timely way have better health outcomes. The number of visits matters only in part; the timing of visits is also crucial, and the insurance program makes that timing more flexible. Furthermore, the initial increase in visits associated with the program may have increased household familiarity and comfort with clinic services. The availability of such health financing instruments and—potentially—explicitly making them available at the point of transfer distribution may be important considerations if countries desire to fully reap health gains from cash transfers.

Our analysis of the health impacts of Tanzania’s CCT program is not without caveats; while self-reported health measures improved, we do not find enduring impacts on children’s anthropometrics. The cash transfers may make children feel better and be more able to carry out daily activities (e.g., attending school and fetching water), but these may not immediately translate into growth—at least not within 2.5 years. However, they suggest a clear increase in reported child well-being.

We find some evidence that overall, health improvements—at least on the intensive margin—are greater in villages with more health workers per capita. In other words, cash transfers can most effectively reduce the number of days individuals are sick when clinics are sufficiently staffed to provide high-quality services. Clear evidence demonstrates that healthier households are likely to have higher incomes, which then drive better health in a virtuous cycle ([Strauss and Thomas, 1998](#)). The evidence from this study demonstrates that supply-side investments (in health care providers) combined with cash transfers (to permit households to insure against health shocks) may be critical catalysts to that virtuous cycle.

References

- Adhvaryu, A. and A. Nyshadham (2015). Returns to treatment in the formal health care sector: Evidence from Tanzania. *American Economic Journal: Economic Policy* 7(3), 29–57.
- Akresh, R., D. de Walque, and H. Kazianga (2014). *Alternative Cash Transfer Delivery Mechanisms: Impacts on Routine Preventative Health Clinic Visits in Burkina Faso*. University of Chicago Press. <http://www.nber.org/chapters/c13377>.
- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association* 103(484).
- Andrews, C., A. Lopez, and J. Baez (2014). What are we learning on safety net impacts? reviewing evidence from 2010-2013. *World Bank Social Protection Discussion Paper Series*.
- Attanasio, O., E. Battistin, E. Fitzsimons, and M. Vera-Hernandez (2005). How effective are conditional cash transfers? evidence from colombia.
- Babbel, B. (2012). Evaluating equity in the provision of primary health care in Tanzania. <https://ir.library.oregonstate.edu/xmlui/handle/1957/32933>.
- Baird, S., J. De Hoop, and B. Özler (2013). Income shocks and adolescent mental health. *Journal of Human Resources* 48(2), 370–403.
- Baird, S., C. McIntosh, and B. Özler (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, qjr032.
- Bank of Tanzania (2015). Interbank foreign exchange market summaries. <https://www.bot-tz.org/FinancialMarkets/IFEMsummaries/IFEMsummaries.asp>.
- Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen (2015). Turning a shove into a nudge? a “labeled cash transfer” for education. *American Economic Journal: Economic Policy* 7(3), 86–125.
- Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.
- Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* 93(3), 491–507.
- Biosca, O. and H. Brown (2014). Boosting health insurance coverage in developing countries: do conditional cash transfer programmes matter in Mexico? *Health policy and planning*, czt109.
- Birn, A.-E. and A. Solórzano (1999). Public health policy paradoxes: science and politics in the Rockefeller Foundation’s hookworm campaign in Mexico in the 1920s. *Social Science & Medicine* 49(9), 1197–1213.
- Bjorkman Nyqvist, M., L. Corno, D. De Walque, and J. Svensson (2015). Using lotteries to incentivize safer sexual behavior: evidence from a randomized controlled trial on HIV prevention. *World Bank Policy Research Working Paper* (7215).
- De Walque, D., W. Dow, and R. Nathan (2014). Rewarding safer sex: conditional cash transfers for HIV/STI prevention. *World Bank Policy Research Working Paper* (7099).
- Dercon, S., J. W. Gunning, and A. Zeitlin (2015). The demand for insurance under limited trust: Evidence from a field experiment in kenya.
- Duflo, E. (2000). Child health and household resources in South Africa: Evidence from the old age pension program. *American Economic Review*, 393–398.

- Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *The World Bank Economic Review* 17(1), 1–25.
- Ekman, B. (2004). Community-based health insurance in low-income countries: a systematic review of the evidence. *Health policy and planning* 19(5), 249–270.
- Evans, D., S. Hausladen, K. Kosec, and N. Reese (2014). *Community-Based Conditional Cash Transfers in Tanzania: Results from a Randomized Trial*. World Bank Publications.
- Evans, D., B. Holtemeyer, and K. Kosec (2016). If you give it, trust will come: The impacts of community-managed cash transfers. *Working Paper*.
- Fiszbein, A. and N. Schady (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications.
- Garcia, M. and C. M. Moore (2012). *The cash dividend: The rise of cash transfer programs in Sub-Saharan Africa*. World Bank Publications.
- Handa, S., A. Peterman, D. Seidenfeld, and G. Tembo (2015). Income transfers and maternal health: Evidence from a national randomized social cash transfer program in Zambia. *Health Economics*. <http://dx.doi.org/10.1002/hec.3136>.
- Haushofer, J. and J. Shapiro (2016). The short-term impact of unconditional cash transfers to the poor: Evidence from Kenya. *Forthcoming, Quarterly Journal of Economics*.
- Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75(4), 800–802.
- Hoddinott, J. (2010). Nutrition and conditional cash transfer programs. In M. Adato and J. Hoddinott (Eds.), *Conditional Cash Transfers in Latin America*. Baltimore: Johns Hopkins University Press.
- Independent Evaluation Group (2011). Social safety nets: An evaluation of world bank support, 2000–2010. Washington, DC: Independent Evaluation Group, the World Bank Group.
- Institute for Health Metrics and Evaluation (2013). The global burden of disease: Generating evidence, guiding policy — Sub-Saharan Africa regional edition. <http://documents.worldbank.org/curated/en/2013/08/18187588/global-burden-disease-generating-evidence-guiding-policy-sub-saharan-africa-regional-edition>.
- Kamuzora, P. and L. Gilson (2007). Factors influencing implementation of the community health fund in Tanzania. *Health Policy and Planning* 22(2), 95–102.
- King, E. M. and J. R. Behrman (2009). Timing and duration of exposure in evaluations of social programs. *The World Bank Research Observer* 24(1), 55–82.
- Leroy, J. L., M. Ruel, and E. Verhofstadt (2009). The impact of conditional cash transfer programmes on child nutrition: a review of evidence using a programme theory framework. *Journal of Development Effectiveness* 1(2), 103–129.
- Levy, D. and J. Ohls (2007). Evaluation of jamaica’s path program: final report. *Report prepared*.
- Macours, K., N. Schady, and R. Vakis (2012). Cash transfers, behavioral changes, and the cognitive development of young children: Evidence from a randomized experiment. *American Economic Journal: Applied Economics* 4(2), 247–273.
- Maluccio, J. and R. Flores (2005). *Impact evaluation of a conditional cash transfer program: The Nicaraguan Red de Protección Social*. International Food Policy Research Institute.
- Marriott, A. (2011). Does health insurance work in Tanzania? <http://www.globalhealthcheck.org/?p=388>.
- Mascarini-Serra, L. et al. (2011). Prevention of soil-transmitted helminth infection. *Journal of global infectious diseases* 3(2), 175.

- Morris, S. S., P. Olinto, R. Flores, E. A. Nilson, and A. C. Figueiro (2004). Conditional cash transfers are associated with a small reduction in the rate of weight gain of preschool children in northeast Brazil. *The Journal of nutrition* 134(9), 2336–2341.
- Mtei, G., J. Mulligan, et al. (2007). Community health funds in Tanzania: A literature review. *Ifakara Health Research and Development Centre, Ifakara*.
- Paxson, C. and N. Schady (2010). Does money matter? the effects of cash transfers on child development in rural Ecuador. *Economic development and cultural change* 59(1), 187–229.
- Powell-Jackson, T. and K. Hanson (2012). Financial incentives for maternal health: impact of a national programme in Nepal. *Journal of health economics* 31(1), 271–284.
- Powell-Jackson, T., S. Mazumdar, and A. Mills (2015). Financial incentives in health: New evidence from India's Janani Suraksha Yojana. *Journal of health economics* 43, 154–169.
- Robertson, L., P. Mushati, J. W. Eaton, L. Dumba, G. Mavise, J. Makoni, C. Schumacher, T. Crea, R. Monasch, L. Sherr, et al. (2013). Effects of unconditional and conditional cash transfers on child health and development in Zimbabwe: a cluster-randomised trial. *The Lancet* 381(9874), 1283–1292.
- Schaffer, M. (2010). xtivreg2: Stata module to perform extended iv/2sls, gmm and ac/hac, liml and k-class regression for panel data models. <http://ideas.repec.org/c/boc/bocode/s456501.html>.
- Strauss, J. and D. Thomas (1998). Health, nutrition, and economic development. *Journal of economic literature*, 766–817.
- World Bank (2013). *Africa's Pulse 8*. http://www.worldbank.org/content/dam/Worldbank/document/Africa/Report/Africas-Pulse-brochure_Vol8.pdf.
- World Health Organization (2014). United Republic of Tanzania: Health Profile. *Africa's Pulse*. <http://www.who.int/gho/countries/tza.pdf?ua=1>.

Table 1: Timeline of CCT program and impact evaluation

Timing	Activity
November 2007 - September 2008	Program design
September - November 2008	Sensitization at regional, district, ward, and community levels
January - May 2009	Baseline survey
September - October 2009	Enrollment of beneficiaries
January 2010	First payments made to beneficiary households
July - September 2011	Midline survey and first round of qualitative data collection
August - October 2012	Endline survey
July - August 2013	Second round of qualitative data collection

Table 2: Effects of treatment on health clinic visits in the last 12 months

	Full sample (1)	0-5 years old (2)	60 and over (3)
Treatment × 2011 (midline)		2.296** (0.872)	1.083*** (0.349)
Treatment × 2012 (endline)	-0.067 (0.253)	-1.042 (0.875)	0.161 (0.344)
2011 (midline)		-3.817*** (0.584)	-1.214*** (0.216)
2012 (endline)	-1.436*** (0.182)	-5.762*** (0.635)	-0.670*** (0.237)
R^2	0.061	0.375	0.018
Baseline mean	2.802	8.272	2.783
Observations	13713	1243	5692

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Midline data are excluded from the full sample because health facility visit data were not collected in the midline survey for those 5-60 years old. Ages refer to age at the time of baseline survey. Fewer refers to those residing in villages in the bottom half of the distribution of baseline health clinic staff per capita, while more refers to those in the top half. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 3: Effects of treatment on compliance with health clinic visits in the last 12 months

	Full sample (1)	0-5 years old (2)	60 and over (3)
Treatment × 2011 (midline)		0.288*** (0.085)	0.259*** (0.031)
Treatment × 2012 (endline)	0.039** (0.016)	0.019 (0.058)	0.086** (0.035)
2011 (midline)		-0.086 (0.067)	-0.120*** (0.026)
2011 (endline)	0.020 (0.012)	0.186*** (0.044)	-0.010 (0.027)
R^2	0.010	0.074	0.027
Observations	13,713	1,243	5,692
Baseline mean	0.831	0.591	0.653

Sources: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). To be in health compliance, those aged 0-5 years must have 6 clinic visits in last 12 months; those 60+ must have 1 clinic visit in the last 12 months. At endline (2012), the condition was loosened from 6 to 2 visits for those aged 2-5 years. Midline data are excluded from the full sample because health facility visit data were not collected in the midline survey for those 5-60 years old. Ages refer to age at the time of baseline survey. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 4: Effects of treatment on take-up of health-related products

	Dummy - owns shoes (1)	Dummy - owns slippers (2)	Insurance expenditures (3)	Dummy - participates in the CHF (4)
Treatment × 2011 (midline)	0.180*** (0.043)	0.054 (0.035)	1.176*** (0.252)	
Treatment × 2012 (endline)	0.179*** (0.047)	0.084** (0.038)	1.516*** (0.284)	0.357*** (0.039)
2011 (midline)	0.129*** (0.028)	0.188*** (0.023)	0.177*** (0.051)	
2012 (endline)	0.126*** (0.031)	0.196*** (0.028)	0.438*** (0.099)	
R^2	0.105	0.107	0.118	0.317
Baseline mean	0.423	0.632	0.181	
Observations	6847	6847	5036	1555

Sources: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Shoe and slipper ownership are individual-level outcomes for those under 18 years old at the time of the baseline survey. Insurance expenditures and CHF participation are household level outcomes. Insurance expenditures refer of total annual medical, car, and life insurance expenditures (thousands TSH). Data on participation in the CHF are only available from the endline survey. Households that report having never heard of the CHF are assumed to not be participating in the CHF. Columns (1) and (2) include individual fixed effects. Column (3) includes household fixed effects. Column (4) includes baseline controls of age, age², sex, and education level of the household head. Also included are dummies for district, household size, having an improved roof, having an improved toilet, having an improved floor, having piped water, village population, the number of years since the CHF began operating in respondent's village, and the first principal component from a PCA using information on ownership of 13 household assets. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 5: Effects of treatment on method used to finance healthcare when addressing main health problem of the last month

	Free treatment (1)	Loan or assistance (2)	Cash or asset (3)	Health insurance (4)
Treatment × 2011 (midline)	0.038 (0.052)	-0.017 (0.034)	-0.176*** (0.062)	0.155*** (0.049)
Treatment × 2012 (endline)	-0.083* (0.048)	-0.013 (0.032)	-0.182*** (0.065)	0.278*** (0.052)
2011 (midline)	-0.007 (0.035)	-0.071** (0.027)	0.054 (0.037)	0.024 (0.015)
2012 (endline)	0.027 (0.037)	-0.074*** (0.023)	0.007 (0.039)	0.040*** (0.013)
R^2	0.009	0.035	0.022	0.142
Baseline mean	0.216	0.103	0.654	0.027
Observations	5365	5365	5365	5365

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Health problem of the last month refers to the last four weeks. Over the 3 rounds of the survey, respondents reported being sick or injured a total of 5,922 times. For 5,409 of these reports, the main treatment financing method was reported. 44 people were excluded from this analysis for reporting financing with either "other" or "differed by provider" since it was not possible to understand how these individuals financed treatment. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 6: Effects of treatment on illness and injury in the last month

	Dummy - ill or injured in last month			Days in last month unable to perform normal daily activities due to illness or injury		
	Full sample (1)	0-5 years old (2)	60 and over (3)	Full sample (4)	0-5 years old (5)	60 and over (6)
Treatment × 2011 (midline)	0.004 (0.026)	-0.011 (0.055)	0.044 (0.040)	-0.210 (0.225)	-0.122 (0.285)	-0.204 (0.489)
Treatment × 2012 (endline)	-0.043 (0.026)	-0.107* (0.063)	-0.002 (0.035)	-0.435* (0.220)	-0.758** (0.358)	-0.353 (0.414)
2011 (midline)	0.002 (0.018)	-0.054* (0.032)	0.032 (0.028)	0.198 (0.165)	-0.206 (0.170)	0.675** (0.323)
2012 (endline)	0.078*** (0.016)	0.031 (0.047)	0.147*** (0.023)	1.076*** (0.147)	0.298 (0.297)	2.389*** (0.269)
R^2	0.006	0.010	0.024	0.012	0.011	0.033
Baseline mean	0.276	0.282	0.388	1.636	1.052	2.786
Observations	20741	1537	5694	20740	1537	5693

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Illness in the last month refers to the last four weeks. Ages refer to age at the time of baseline survey. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 7: Effects of treatment on type of health provider visited to address main health problem of the last month

	Public				Private			
	None	Dispensary	District, region, or referral hospital	Health center	Pharmacy or chemist	Healer, herbalist, or faith healer	Dispensary, hospital, clinic, or store	Mission dispensary or hospital
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment × 2011 (midline)	-0.119*** (0.041)	0.174*** (0.048)	-0.036 (0.024)	0.001 (0.025)	-0.035 (0.040)	0.024 (0.016)	-0.007 (0.017)	-0.002 (0.008)
Treatment × 2012 (endline)	-0.054 (0.037)	0.148*** (0.050)	-0.003 (0.029)	-0.021 (0.026)	-0.055 (0.047)	-0.010 (0.022)	-0.001 (0.019)	-0.004 (0.014)
2011 (midline)	0.100*** (0.032)	-0.033 (0.035)	0.010 (0.019)	0.030 (0.020)	-0.066** (0.034)	-0.027* (0.014)	-0.011 (0.012)	-0.003 (0.004)
2012 (endline)	0.005 (0.027)	0.053 (0.041)	0.017 (0.025)	0.029 (0.020)	-0.088** (0.041)	-0.012 (0.017)	-0.004 (0.015)	0.001 (0.013)
R^2	0.016	0.029	0.005	0.006	0.024	0.005	0.002	0.001
Baseline mean	0.154	0.395	0.059	0.035	0.274	0.047	0.023	0.013
Observations	5889	5889	5889	5889	5889	5889	5889	5889

Sources: Authors' calculations based on 2009, 2011, and 2012 household survey data. (2012).

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Health problem of the last month refers to the last four weeks. Over the 3 rounds of the survey, respondents reported being sick or injured a total of 5,922 times. In all of those reports, the most important health provider was reported. 33 people were excluded from this analysis for reporting that the most important health provider was other. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table 8: Heterogeneous treatment effects by staff/capita

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Effects of treatment on health clinic visits in the last 12 months</i>						
	Full sample	0-5 years old	60 and over			
T × 2011 × fewer		2.975*** (1.036)	0.875* (0.477)			
T × 2011 × more		1.426 (1.509)	1.279** (0.511)			
T × 2012 × fewer	-0.001 (0.373)	-1.483 (1.114)	0.199 (0.511)			
T × 2012 × more	-0.144 (0.331)	-0.234 (1.223)	0.126 (0.466)			
Observations	13713	1243	5692			
2011 p-value of difference		0.400	0.565			
2012 p-value of difference	0.775	0.453	0.916			
Baseline mean (fewer)	2.598	7.553	2.501			
Baseline mean (more)	3.004	9.104	3.051			
<i>Panel B: Effects of treatment on illness and injury in the last month</i>						
		Dummy - ill or injured in last month		Days in last month unable to perform normal daily activities due to illness or injury		
	Full sample	0-5 years old	60 and over	Full sample	0-5 years old	60 and over
T × 2011 × fewer	0.016 (0.039)	-0.011 (0.064)	0.014 (0.060)	0.083 (0.317)	-0.117 (0.332)	0.219 (0.749)
T × 2011 × more	-0.006 (0.032)	-0.021 (0.084)	0.072 (0.055)	-0.493 (0.309)	-0.186 (0.443)	-0.580 (0.621)
T × 2012 × fewer	-0.029 (0.040)	-0.110 (0.084)	0.008 (0.047)	0.071 (0.307)	-0.486 (0.387)	0.772 (0.527)
T × 2012 × more	-0.056* (0.031)	-0.112 (0.093)	-0.019 (0.048)	-0.959*** (0.281)	-1.135* (0.655)	-1.505*** (0.553)
Observations	20741	1537	5694	20740	1537	5693
2011 p-value of difference	0.668	0.926	0.474	0.197	0.901	0.414
2012 p-value of difference	0.581	0.990	0.687	0.016	0.397	0.004
Baseline mean (fewer)	0.284	0.297	0.380	1.662	1.081	2.763
Baseline mean (more)	0.267	0.264	0.396	1.610	1.017	2.809

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Midline and endline treatment effects are abbreviated T × 2011 and T × 2012, respectively. Fewer refers to those residing in villages in the bottom half of the distribution of baseline health clinic staff per capita, while more refers to those in the top half. Ages refer to age at the time of baseline survey. In panel A midline data are excluded from the full sample because health facility visit data were not collected in the midline survey for those 5-60 years old. In panel B illness in the last month refers to the last four weeks. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.10.

Table 9: Heterogeneous treatment effects by degree of poverty

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Effects of treatment on type of health provider visited to address main health problem of the last month</i>								
	Public				Private			
	None	Dispensary	District, region, or referral hospital	Health center	Pharmacy or chemist	Healer, herbalist, or faith healer	Dispensary, hospital, clinic, or store	Mission dispensary or hospital
T × 2011 × extremely poor	-0.100* (0.058)	0.213*** (0.072)	-0.046 (0.031)	-0.054 (0.034)	0.026 (0.066)	0.001 (0.024)	-0.044** (0.022)	0.005 (0.014)
T × 2011 × moderately poor	-0.136*** (0.050)	0.135* (0.073)	-0.031 (0.035)	0.062* (0.036)	-0.103* (0.058)	0.051** (0.025)	0.031 (0.025)	-0.010 (0.015)
T × 2012 × extremely poor	-0.017 (0.050)	0.166*** (0.061)	-0.043 (0.038)	-0.031 (0.032)	-0.024 (0.063)	-0.022 (0.031)	-0.034 (0.022)	0.006 (0.022)
T × 2012 × moderately poor	-0.091* (0.051)	0.128* (0.070)	0.035 (0.037)	-0.007 (0.033)	-0.090 (0.069)	0.005 (0.035)	0.035 (0.030)	-0.014 (0.015)
Observations	5889	5889	5889	5889	5889	5889	5889	5889
2011 p-value of difference	0.604	0.475	0.744	0.020	0.185	0.163	0.018	0.546
2012 p-value of difference	0.294	0.661	0.117	0.541	0.487	0.575	0.058	0.404
Baseline mean (extremely poor)	0.190	0.381	0.029	0.044	0.268	0.060	0.016	0.012
Baseline mean (moderately poor)	0.120	0.409	0.087	0.028	0.279	0.036	0.029	0.013
<i>Panel B: Effects of treatment on take-up of health-related products</i>								
	Dummy - owns shoes	Dummy - owns slippers	Insurance expenditures	Dummy - CHF participation	<i>Panel C: Effects of treatment on method used to finance healthcare</i>			
					Free treatment	Loan or assistance	Cash or asset	Health insurance
T × 2011 × extremely poor	0.300*** (0.065)	0.095* (0.051)	1.398*** (0.291)		0.055 (0.081)	-0.070 (0.046)	-0.146 (0.097)	0.161** (0.068)
T × 2011 × moderately poor	0.079* (0.044)	0.015 (0.035)	0.871*** (0.254)		0.023 (0.053)	0.032 (0.045)	-0.206*** (0.071)	0.150*** (0.050)
T × 2012 × extremely poor	0.246*** (0.067)	0.159*** (0.049)	1.753*** (0.308)	0.380*** (0.050)	-0.116 (0.073)	-0.071* (0.041)	-0.148 (0.096)	0.335*** (0.072)
T × 2012 × moderately poor	0.120** (0.055)	0.009 (0.046)	1.219*** (0.378)	0.328*** (0.040)	-0.054 (0.057)	0.042 (0.044)	-0.215*** (0.070)	0.227*** (0.047)
Observations	6847	6847	5036	1555	5365	5365	5365	5365
2011 p-value of difference	0.002	0.138	0.059		0.717	0.095	0.612	0.883
2012 p-value of difference	0.106	0.011	0.174	0.295	0.489	0.051	0.534	0.124
Baseline mean (extremely poor)	0.292	0.571	0.093		0.247	0.128	0.610	0.015
Baseline mean (moderately poor)	0.525	0.680	0.289		0.188	0.081	0.693	0.038

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Midline and endline treatment effects are abbreviated T × 2011 and T × 2012, respectively. Degree of poverty refers to the value at the time of the baseline survey on an index of asset ownership. The index is the first principal component from a PCA using information on ownership of 13 household assets. Extremely poor refers to those in the bottom half, while moderately poor refers to those in the top half. Standard errors are in parentheses and clustered at the village level. *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.10.

Panel A: Health problem of the last month refers to the last four weeks. Over the 3 rounds of the survey, respondents reported being sick or injured a total of 5,922 times. In all of those reports, the most important health provider was reported. 33 people were excluded from this analysis for reporting that the most important health provider was other. All specifications include individual fixed effects.

Panel B: Shoe and slipper ownership are individual-level outcomes for those at least 18 years old at the time of the baseline survey. Insurance expenditures and CHF participation are household level outcomes. Insurance expenditures refer to total annual medical, car, and life insurance expenditures (thousands TSH). Data on participation in the CHF are only available from the endline survey. Households that report having never heard of the CHF are assumed to not be participating in the CHF. Degree of poverty refers to the value at the time of the baseline survey on an index of asset ownership. The index is the first principal component from a PCA using information on ownership of 13 household assets. Columns (1) and (2) include individual fixed effects. Column (3) includes household fixed effects. Column (4) includes baseline controls of age, age², sex, and education level of the household head. Also included are dummies for district, household size, having an improved roof, having an improved toilet, having an improved floor, having piped water, village population, the number of years since the CHF began operating in respondent's village, and the asset index used to separate moderately and extreme poverty.

Panel C: Method used to finance healthcare for the main health problem of the last four weeks. Over the 3 rounds of the survey, respondents reported being sick or injured a total of 5,922 times. In 5,409 of those 5,922 reports, the main treatment financing method was reported. 44 people were excluded from this analysis for reporting financing with either "other" or "differed by provider" since it was not possible to understand how these individuals financed treatment. All specifications include individual fixed effects.

Table A1: Baseline (2009) means of outcomes and key characteristics by treatment assignment

	Treatment (T)		Control (C)		Difference (T-C)	
	Mean	N	Mean	N	Mean	S.E.
<i>Panel A: Outcome variables</i>						
Clinic Visits and Health, Full Sample						
# health clinic visits in last 12 months	2.83	3462	2.77	3456	0.06	(0.26)
Dummy - ill or injured in last month	0.29	3462	0.26	3456	0.02	(0.02)
# of days unable to do normal activities	1.68	3462	1.59	3455	0.08	(0.14)
Clinic Visits and Health, 60 and over						
# health clinic visits in last 12 months	2.91	1049	2.67	1160	0.24	(0.35)
Dummy - ill or injured in last month	0.39	1049	0.38	1160	0.01	(0.03)
# of days unable to do normal activities	2.78	1049	2.79	1159	-0.02	(0.31)
Clinic Visits and Health, 0-5 years old						
# health clinic visits in last 12 months	8.21	309	8.33	312	-0.12	(0.70)
Dummy - ill or injured in last month	0.31	309	0.25	312	0.06*	(0.04)
# of days unable to do normal activities	1.31	309	0.80	312	0.51**	(0.21)
Health-related products						
Dummy - owns shoes	0.38	1515	0.47	1441	-0.10**	(0.05)
Dummy - owns slippers	0.62	1515	0.65	1441	-0.03	(0.03)
Insurance expenditures, thousands Tsh	0.26	881	0.11	879	0.15*	(0.08)
Specific activities of daily living, 60 and over						
Dummy - can do vigorous activity	0.35	1049	0.36	1160	-0.02	(0.04)
Dummy - can walk uphill	0.77	1049	0.75	1160	0.01	(0.03)
Dummy - can bend over or stoop	0.97	1049	0.96	1160	0.01	(0.01)
Dummy - can walk over 1km	0.86	1049	0.85	1160	0.01	(0.02)
Dummy - can walk over 100m	0.97	1049	0.96	1160	0.01	(0.01)
Dummy - can use bath or toilet	0.98	1049	0.97	1160	0.00	(0.01)
Ordinary activities index	4.89	1049	4.86	1160	0.03	(0.09)
Anthropometrics						
Height-for-age z-score	-1.46	231	-1.25	240	-0.21	(0.14)
Weight-for-age z-score	-0.90	208	-0.72	189	-0.18*	(0.10)
Weight-for-height z-score	0.06	187	0.04	176	0.02	(0.11)
BMI-for-age z-score	0.23	187	0.16	177	0.07	(0.12)
Height (cm)	87.38	234	87.10	241	0.28	(1.14)
Weight (kg)	12.22	253	12.07	253	0.14	(0.26)
MUAC (mm)	155.68	230	155.83	232	-0.15	(1.42)
Healthcare location						
None	0.17	993	0.13	907	0.04	(0.03)
Public: dispensary	0.38	993	0.41	907	-0.02	(0.04)
Public: district, region, or referral hospital	0.05	993	0.07	907	-0.02	(0.02)
Public: health center	0.03	993	0.04	907	0.00	(0.02)
Private: Pharmacy or chemist	0.28	993	0.27	907	0.01	(0.03)
Private: Healer, herbalist, or faith healer	0.05	993	0.05	907	0.00	(0.01)
Private: dispensary, hospital, clinic, or store	0.02	993	0.02	907	0.00	(0.01)
Private: Mission dispensary or hospital	0.01	993	0.01	907	0.00	(0.01)
Household size						
Total	3.94	879	3.94	878	0.00	(0.21)
0-5 years	0.35	879	0.36	878	0.00	(0.04)
60 and over	1.19	879	1.32	878	-0.13**	(0.05)
Healthcare financing method						
Free treatment	0.20	885	0.23	826	-0.03	(0.03)
Loan or assistance	0.11	885	0.10	826	0.01	(0.02)
Cash or asset	0.65	885	0.66	826	0.00	(0.03)
Health insurance	0.04	885	0.02	826	0.02	(0.02)

table continued on next page...

	Treatment (T)		Control (C)		Difference (T-C)	
	Mean	N	Mean	N	Mean	S.E.
<i>Panel B: Individual and household characteristics</i>						
Individual characteristics						
Age	35.54	3462	37.04	3456	-1.49	(1.20)
Dummy - male	0.47	3462	0.45	3456	0.02	(0.01)
Dummy - has less than Standard 1 education	0.53	3459	0.54	3451	0.00	(0.02)
Dummy - has Standard 1-4 education	0.22	3459	0.22	3451	0.00	(0.01)
Dummy - has at least Standard 5 education	0.24	3459	0.24	3451	0.00	(0.02)
Dummy - literate	0.41	3462	0.42	3456	-0.01	(0.03)
Household characteristics						
Dummy - household has improved roof	0.33	880	0.37	878	-0.04	(0.06)
Dummy - household has improved floor	0.03	880	0.09	878	-0.06**	(0.02)
Dummy - household has toilet facilities	0.69	880	0.76	879	-0.07	(0.04)
Dummy - household has piped water	0.30	880	0.32	879	-0.01	(0.08)
Dummy - head of household is male	0.63	879	0.59	878	0.04	(0.03)
Village characteristics						
Health staff per 1000 villagers (2009)	3.09	40	2.84	40	0.24	(0.94)

Source: Authors' calculations based on baseline (2009) household survey data.

Notes: Treatment indicates assignment to treatment. Illness in the last month refers to the last four weeks. The universe of individuals used to summarize the healthcare location and healthcare financing method outcomes is individuals who reported being ill or injured in the last month (four weeks). Ordinary activities index is the sum of the six activity dummies; its range is 0 to 6. Specific activities of daily living are summarized for those at least 60 years old because data were unavailable for individuals under 60 years old at the time of the midline and endline surveys. Shoe and slipper ownership are summarized for those 18 years old or younger in the baseline survey. Insurance expenditures is a household level outcome, and it refers to total annual medical, car, and life insurance expenditures. BMI is body mass index and MUAC is middle upper-arm circumference. Ages refer to age at time of baseline survey. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A2: Attrition after baseline survey

	Household is in ... survey				Individual is in ... survey				Individual 0-5 or 60+ is in ... survey			
	midline		endline		midline		endline		midline		endline	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment village	0.002 (0.016)	-0.147 (0.178)	0.023 (0.020)	0.132 (0.247)	0.001 (0.017)	-0.096 (0.176)	0.031 (0.021)	-0.145 (0.244)	0.000 (0.018)	-0.163 (0.234)	0.021 (0.023)	-0.347 (0.341)
Dummy - household head male		0.029 (0.020)		0.040* (0.022)		0.045* (0.024)		0.042* (0.024)		0.019 (0.027)		0.041 (0.034)
Dummy - household head male \times Treatment		0.009 (0.028)		0.018 (0.034)		-0.026 (0.034)		0.038 (0.036)		0.011 (0.045)		0.036 (0.053)
Household head age		0.002 (0.003)		0.012** (0.006)		0.005 (0.003)		0.005 (0.006)		-0.003 (0.006)		-0.007 (0.008)
Household head age \times Treatment		0.004 (0.006)		-0.003 (0.008)		0.004 (0.005)		0.007 (0.008)		0.009 (0.008)		0.016 (0.010)
Household head age ²		-0.000 (0.000)		-0.000** (0.000)		-0.000* (0.000)		-0.000 (0.000)		0.000 (0.000)		0.000 (0.000)
Household head age ² \times Treatment		-0.000 (0.000)		0.000 (0.000)		-0.000 (0.000)		-0.000 (0.000)		-0.000 (0.000)		-0.000* (0.000)
Head has some education		-0.013 (0.021)		-0.026 (0.028)		0.043 (0.027)		0.034 (0.032)		0.022 (0.030)		-0.033 (0.040)
Head has some education \times Treatment		0.032 (0.030)		0.003 (0.038)		0.003 (0.035)		0.028 (0.042)		-0.011 (0.050)		0.077 (0.062)
Asset index		0.011*** (0.004)		0.002 (0.008)		-0.009** (0.004)		-0.014** (0.006)		-0.008* (0.004)		-0.016** (0.007)
Asset index \times Treatment		-0.012 (0.007)		-0.001 (0.012)		0.002 (0.006)		-0.004 (0.010)		-0.001 (0.007)		-0.001 (0.011)
Dummy - male						-0.018 (0.014)		0.007 (0.016)		0.003 (0.018)		0.021 (0.023)
Dummy - male \times Treatment						0.006 (0.020)		-0.017 (0.023)		-0.025 (0.034)		-0.067* (0.037)
Age						0.002 (0.002)		0.004*** (0.001)		0.013*** (0.002)		0.015*** (0.002)
Age \times Treatment						-0.000 (0.002)		-0.001 (0.002)		-0.005 (0.003)		-0.007** (0.003)
Age ²						-0.000 (0.000)		-0.000** (0.000)		-0.000*** (0.000)		-0.000*** (0.000)
Age ² \times Treatment						0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		0.000** (0.000)
Some education						-0.042** (0.018)		-0.062*** (0.021)		-0.013 (0.029)		0.030 (0.036)
Some education \times Treatment						-0.027 (0.025)		-0.038 (0.029)		0.034 (0.054)		-0.034 (0.065)
Constant	0.913*** (0.013)	0.890*** (0.106)	0.857*** (0.015)	0.546*** (0.181)	0.780*** (0.013)	0.636*** (0.120)	0.686*** (0.017)	0.506*** (0.185)	0.818*** (0.014)	0.757*** (0.183)	0.732*** (0.017)	0.785*** (0.257)
Observations	1,764	1,757	1,764	1,757	6,918	6,910	6,918	6,910	3,016	3,010	3,016	3,010
R-squared	0.000	0.014	0.001	0.016	0.000	0.019	0.001	0.030	0.000	0.040	0.001	0.047
Test of joint signif. of interactions (p-value)		0.677		0.889		0.921		0.601		0.902		0.119

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Those aged 0-5 years and 60+ are those for whom program health conditions applied. The asset index is the first principal component from a PCA using information on ownership of 13 household assets. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A3: Effects of treatment (on the treated) on health clinic visits in the last 12 months

	Full sample: baseline and endline (1)	0-5 years old (2)	60 and over (3)
<i>Panel A: Effect of treatment on the treated</i>			
Treatment × 2011 (midline)		2.637*** (0.966)	1.138*** (0.361)
Treatment × 2012 (endline)	-0.072 (0.268)	-1.153 (0.954)	0.170 (0.359)
2011 (midline)		-3.812*** (0.580)	-1.217*** (0.215)
2012 (endline)	-1.436*** (0.181)	-5.751*** (0.632)	-0.671*** (0.236)
R^2	0.061	0.374	0.018
Baseline mean	2.802	8.272	2.783
Observations	9622	1092	5339
<i>Panel B: Heterogeneous treatment effects by staff/capita</i>			
Treatment effect for those with fewer (endline)	-0.001 (0.389)	-1.637 (1.212)	0.207 (0.527)
Treatment effect for those with more (endline)	-0.157 (0.359)	-0.259 (1.348)	0.135 (0.493)
Treatment effect for those with fewer (midline)		3.388*** (1.126)	0.908* (0.488)
Treatment effect for those with more (midline)		1.650 (1.706)	1.366** (0.536)
p-value of difference (midline)		0.395	0.527
p-value of difference (endline)	0.768	0.447	0.920
Baseline mean for those with fewer	2.598	7.553	2.501
Baseline mean for those with more	3.004	9.104	3.051

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Midline data are excluded from the full sample because health facility visit data were not collected in the midline survey for those 5-60 years old. Ages refer to age at the time of baseline survey. Fewer refers to those residing in villages in the bottom half of the distribution of baseline health clinic staff per capita, while more refers to those in the top half. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A4: Effects of treatment (on the treated) on take-up of health-related products

	Dummy - owns shoes (1)	Dummy - owns slippers (2)	Insurance expenditures (thousands Tsh) (3)	Dummy - participates in the CHF (4)
<i>Panel A: Effect of treatment on the treated</i>				
Treatment × 2011 (midline)	0.194*** (0.047)	0.058 (0.037)	1.262*** (0.267)	
Treatment × 2012 (endline)	0.192*** (0.050)	0.089** (0.041)	1.623*** (0.302)	0.379*** (0.040)
2011 (midline)	0.128*** (0.028)	0.188*** (0.023)	0.172*** (0.051)	
2012 (endline)	0.125*** (0.031)	0.195*** (0.028)	0.433*** (0.099)	
R^2	0.106	0.106	0.124	0.313
Baseline mean	0.423	0.632	0.181	
Observations	6138	6138	4920	1555
<i>Panel B: Heterogeneous treatment effects by degree of poverty</i>				
Treatment effect for extremely poor (midline)	0.318*** (0.069)	0.101* (0.054)	1.469*** (0.302)	
Treatment effect for moderately poor (midline)	0.086* (0.048)	0.016 (0.038)	0.962*** (0.281)	
Treatment effect for extremely poor (endline)	0.262*** (0.070)	0.170*** (0.052)	1.842*** (0.319)	0.399*** (0.051)
Treatment effect for moderately poor (endline)	0.129** (0.058)	0.009 (0.049)	1.337*** (0.412)	0.355*** (0.043)
p-value of difference (midline)	0.002	0.137	0.084	
p-value of difference (endline)	0.099	0.009	0.226	0.409
Baseline mean for extremely poor	0.292	0.571	0.093	
Baseline mean for moderately poor	0.525	0.680	0.289	

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Shoe and slipper ownership are individual-level outcomes for those at least 18 years old at the time of the baseline survey. Insurance expenditures and CHF participation are household level outcomes. Insurance expenditures refer to total annual medical, car, and life insurance expenditures. Data on participation in the CHF are only available from the endline survey. Households that report having never heard of the CHF are assumed to not be participating in the CHF. Degree of poverty refers to the value at the time of the baseline survey on an index of asset ownership. The index is the first principal component from a PCA using information on ownership of 13 household assets. Extremely poor refers to those in the bottom half, while moderately poor refers to those in the top half. Columns (1) and (2) include individual fixed effects. Column (3) includes household fixed effects. Column (4) includes baseline controls of age, age², sex, and education level of the household head. Also included are dummies for district, household size, having an improved roof, having an improved toilet, having an improved floor, having piped water, village population, the number of years since the CHF began operating in respondent's village, and the asset index used to separate moderately and extreme poverty. Standard errors are in parentheses and clustered at the village level. *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.10.

Table A5: Attributes predicting whether an individual was ill or injured in the last month

	Full sample				0-5 or 60+ years			
	midline		endline		midline		endline	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment village	0.033 (0.020)	0.179 (0.152)	-0.013 (0.022)	0.101 (0.196)	0.062** (0.031)	0.253 (0.369)	-0.006 (0.034)	0.054 (0.457)
Dummy - household head male		-0.020 (0.022)		0.028 (0.019)		-0.101** (0.039)		0.018 (0.040)
Dummy - household head male \times Treatment		-0.025 (0.029)		-0.002 (0.031)		0.030 (0.055)		0.056 (0.061)
Household head age		0.003 (0.003)		-0.000 (0.005)		-0.002 (0.009)		-0.003 (0.012)
Household head age \times Treatment		-0.005 (0.005)		-0.003 (0.006)		-0.003 (0.012)		0.001 (0.014)
Household head age ²		-0.000 (0.000)		-0.000 (0.000)		0.000 (0.000)		0.000 (0.000)
Household head age ² \times Treatment		0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		-0.000 (0.000)
Head has some education		0.002 (0.022)		-0.034 (0.028)		0.009 (0.045)		-0.011 (0.052)
Head has some education \times Treatment		0.024 (0.032)		0.033 (0.038)		-0.015 (0.060)		0.013 (0.066)
Asset index		0.002 (0.004)		-0.007 (0.005)		0.006 (0.009)		-0.010 (0.011)
Asset index \times Treatment		-0.013** (0.006)		0.019** (0.008)		-0.024* (0.014)		0.010 (0.016)
Dummy - male		-0.002 (0.014)		-0.045*** (0.016)		0.043 (0.033)		-0.024 (0.038)
Dummy - male \times Treatment		-0.035* (0.020)		0.000 (0.021)		-0.068 (0.047)		-0.064 (0.054)
Age		-0.003* (0.001)		-0.002 (0.002)		0.003 (0.003)		0.003 (0.004)
Age \times Treatment		0.002 (0.002)		0.000 (0.002)		-0.003 (0.004)		-0.001 (0.005)
Age ²		0.000*** (0.000)		0.000*** (0.000)		-0.000 (0.000)		0.000 (0.000)
Age ² \times Treatment		-0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		0.000 (0.000)
Some education		-0.014 (0.018)		-0.000 (0.023)		0.013 (0.042)		0.023 (0.047)
Some education \times Treatment		-0.033 (0.024)		-0.035 (0.030)		-0.033 (0.060)		-0.008 (0.065)
Constant	0.247*** (0.016)	0.161 (0.110)	0.325*** (0.015)	0.322** (0.141)	0.361*** (0.023)	0.251 (0.270)	0.473*** (0.023)	0.311 (0.384)
Observations	6,985	6,983	6,838	6,836	2,300	2,300	2,101	2,101
R-squared	0.001	0.073	0.000	0.085	0.004	0.051	0.000	0.059
Test of joint signif. of interactions (p-value)		0.060		0.241		0.369		0.612

Source: Authors' calculations based on baseline 2009, 2011, and 2012 household survey data.

Notes: Treatment indicates assignment to treatment. The asset index is the first principal component from a PCA using information on ownership of 13 household assets. P-values reported are the result of testing for the joint significance of all of the interaction terms. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A6: Effects of treatment (on the treated) on illness and injury in the last month

	Dummy - ill or injured in last month			Days in last month unable to perform normal daily activities due to illness or injury		
	Full sample (1)	0-5 years old (2)	60 and over (3)	Full sample (4)	0-5 years old (5)	60 and over (6)
<i>Panel A: Effect of treatment on the treated</i>						
Treatment × 2011 (midline)	0.004 (0.027)	-0.012 (0.060)	0.046 (0.042)	-0.225 (0.240)	-0.135 (0.311)	-0.215 (0.511)
Treatment × 2012 (endline)	-0.046* (0.027)	-0.118* (0.068)	-0.002 (0.036)	-0.465** (0.234)	-0.835** (0.390)	-0.371 (0.433)
2011 (midline)	0.002 (0.018)	-0.054* (0.032)	0.032 (0.028)	0.199 (0.165)	-0.205 (0.171)	0.675** (0.322)
2012 (endline)	0.078*** (0.015)	0.032 (0.047)	0.147*** (0.023)	1.077*** (0.147)	0.306 (0.297)	2.390*** (0.268)
R^2	0.007	0.013	0.024	0.012	0.010	0.033
Baseline mean	0.276	0.282	0.388	1.636	1.052	2.786
Observations	18180	1431	5341	18180	1431	5341
<i>Panel B: Heterogeneous treatment effects by staff/capita</i>						
Treatment effect for those with fewer (midline)	0.016 (0.040)	-0.012 (0.069)	0.014 (0.062)	0.087 (0.331)	-0.129 (0.355)	0.227 (0.771)
Treatment effect for those with more (midline)	-0.007 (0.035)	-0.023 (0.094)	0.077 (0.058)	-0.539 (0.337)	-0.204 (0.494)	-0.619 (0.659)
Treatment effect for those with fewer (endline)	-0.030 (0.041)	-0.120 (0.090)	0.009 (0.049)	0.075 (0.321)	-0.532 (0.418)	0.800 (0.543)
Treatment effect for those with more (endline)	-0.061* (0.033)	-0.124 (0.102)	-0.020 (0.051)	-1.045*** (0.306)	-1.262* (0.721)	-1.604*** (0.590)
p-value of difference (midline)	0.667	0.928	0.458	0.185	0.901	0.404
p-value of difference (endline)	0.553	0.979	0.683	0.012	0.381	0.003
Baseline mean for those with fewer	0.284	0.297	0.380	1.662	1.081	2.763
Baseline mean for those with more	0.267	0.264	0.396	1.610	1.017	2.809

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Illness in the last month refers to the last four weeks. Ages refer to age at the time of baseline survey. Fewer refers to those residing in villages in the bottom half of the distribution of baseline health clinic staff per capita, while more refers to those in the top half. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A7: Effects of treatment on mortality

	Full sample (1)	0-5 years old (2)	60 and over (3)
Treatment	0.005 (0.003)	-0.003 (0.006)	0.010 (0.009)
2012 (endline)	-0.025*** (0.003)	-0.010** (0.005)	-0.066*** (0.007)
R^2	0.065	0.032	0.066
Observations	13042	1192	3993

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: The analysis includes (up to) two observations per individual: a midline observation (which examines deaths between baseline and midline) and an endline observation (which examines deaths between midline and endline). At midline, the outcome is a dummy for the individual being dead at midline (individuals not in sample at baseline take on a missing value). At endline, the outcome is a dummy for the individual being dead at endline (individuals not in sample at midline take on a missing value). Baseline controls not shown, include the age, age², sex, and education level of both the respondent and the household head. Also included are district fixed effects and dummies for gender, household size, having an improved roof, having an improved toilet, having an improved floor, having piped water, village population, and the first principal components from a PCA using information on ownership of 13 household assets at baseline. Ages refer to age at the time of baseline survey. Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A8: Effects of treatment on household size

	Full sample (1)	0-5 years old (2)	60 and over (3)
Treatment \times 2011 (midline)	0.038 (0.103)	0.021 (0.035)	0.028 (0.029)
Treatment \times 2012 (endline)	-0.017 (0.123)	0.017 (0.044)	0.013 (0.035)
2011 (midline)	0.035 (0.079)	-0.073*** (0.024)	-0.074*** (0.021)
2012 (endline)	0.134 (0.104)	-0.065** (0.031)	-0.040 (0.025)
R^2	0.002	0.005	0.005
Baseline mean	3.937	0.353	1.257
Observations	5028	5028	5028

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Ages refer to age at the time of baseline survey. All specifications include household fixed effects. Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A9: Robustness of results to Poisson estimation for count data outcomes

estimate	universe	column	OLS		Poisson			
			S.E. clustered at village level		Robust S.E.		Bootstrap S.E.	
			Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
			(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Effects of treatment on health clinic visits of the treated in the last 12 months (table 2)</i>								
Treatment × 2011	Total	1	.	(.)	.	(.)	.	(.)
Treatment × 2012	Total	1	-0.067	(0.253)	0.003	(0.072)	0.003	(0.072)
Treatment × 2011	0-5 years	2	2.296***	(0.872)	0.340***	(0.122)	0.340***	(0.126)
Treatment × 2012	0-5 years	2	-1.042	(0.875)	-0.348*	(0.188)	-0.348**	(0.174)
Treatment × 2011	60 and over	3	1.083***	(0.349)	0.551***	(0.096)	0.551***	(0.097)
Treatment × 2012	60 and over	3	0.161	(0.344)	0.092	(0.106)	0.092	(0.109)
<i>Panel B: Effects of treatment on number of days unable to perform normal daily activities (table 6)</i>								
Treatment × 2011	Total	1	-0.210	(0.225)	-0.132	(0.094)	-0.132	(0.097)
Treatment × 2012	Total	1	-0.435*	(0.220)	-0.222**	(0.092)	-0.222**	(0.109)
Treatment × 2011	0-5 years	2	-0.122	(0.285)	-0.006	(0.311)	-0.006	(0.294)
Treatment × 2012	0-5 years	2	-0.758**	(0.358)	-0.747**	(0.314)	-0.747**	(0.346)
Treatment × 2011	60 and over	3	-0.204	(0.489)	-0.087	(0.131)	-0.087	(0.145)
Treatment × 2012	60 and over	3	-0.353	(0.414)	-0.114	(0.127)	-0.114	(0.134)
<i>Panel C: Effects of treatment on household size (table A8)</i>								
Treatment × 2011	Total	1	0.038	(0.103)	0.009	(0.023)	0.009	(0.023)
Treatment × 2012	Total	1	-0.017	(0.123)	-0.004	(0.028)	-0.004	(0.032)
Treatment × 2011	0-5 years	2	0.021	(0.035)	0.058	(0.096)	0.058	(0.101)
Treatment × 2012	0-5 years	2	0.017	(0.044)	0.048	(0.108)	0.048	(0.123)
Treatment × 2011	60 and over	3	0.028	(0.029)	0.018	(0.021)	0.018	(0.020)
Treatment × 2012	60 and over	3	0.013	(0.035)	0.008	(0.025)	0.008	(0.023)

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Standard errors in column 6 are bootstrapped over 100 samples with replacement. The first row of Panel A has missing data since health clinic visit data for those aged 5 - 60 were not collected at midline. Column refers to the column in which the estimate appears in the original table. All specifications include individual fixed effects. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A10: Effects of treatment on ability to perform ordinary activities

	dummy - can ...						Ordinary activities index
	do vigorous activity	walk uphill	bend over or stoop	walk more than 1km	walk more than 100m	use bath or toilet	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment × 2011 (midline)	0.026 (0.049)	-0.012 (0.039)	-0.004 (0.009)	-0.007 (0.023)	-0.017 (0.010)	0.002 (0.007)	0.003 (0.101)
Treatment × 2012 (endline)	0.013 (0.059)	-0.022 (0.035)	0.012 (0.014)	-0.026 (0.023)	-0.015** (0.006)	0.013 (0.011)	-0.048 (0.099)
2011 (midline)	0.188*** (0.036)	0.078*** (0.025)	-0.005 (0.006)	0.028* (0.015)	0.003 (0.008)	-0.001 (0.004)	0.289*** (0.066)
2012 (endline)	-0.177*** (0.042)	-0.027 (0.026)	-0.046*** (0.010)	-0.049*** (0.018)	0.011** (0.005)	-0.027*** (0.009)	-0.050 (0.070)
R^2	0.156	0.027	0.018	0.025	0.004	0.009	0.059
Baseline mean	0.356	0.760	0.968	0.855	0.962	0.974	4.875
Observations	5685	5685	5685	5685	5403	5685	5403

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Activity index is the sum of the six activity dummies. Only those at least 60 years old at the time of the baseline are included, due to data availability. All specifications include individual fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A11: Effects of treatment on anthropometrics for children aged 0-5

	z-scores						
	Height-for-age z-score	Weight-for-age z-score	Weight-for-height z-score	BMI-for-age z-score	Height (cm)	Weight (kg)	MUAC (mm)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment × 2011 (midline)	0.105 (0.241)	-0.022 (0.208)	-0.281 (0.303)	-0.351 (0.315)	-0.724 (1.191)	-0.179 (0.228)	-1.422 (2.370)
Treatment × 2012 (endline)	0.222 (0.363)	-0.113 (0.216)	-0.425 (0.332)	-0.488 (0.371)	-0.671 (1.287)	0.008 (0.262)	0.112 (2.650)
2011 (midline)	0.001 (0.183)	0.228 (0.139)	0.501*** (0.191)	0.497** (0.212)	-0.115 (0.492)	0.553*** (0.156)	-0.615 (1.503)
2012 (endline)	0.610** (0.262)	0.805*** (0.142)	0.552*** (0.207)	0.474** (0.240)	0.259 (0.834)	0.717*** (0.180)	0.092 (1.624)
R^2	0.081	0.087	0.075	0.073	0.065	0.074	0.037
Baseline mean	-1.354	-0.812	0.052	0.197	87.240	12.146	155.753
Observations	1184	1204	1079	1073	1240	1403	1234

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Treatment estimates are estimates of the effect of living in a treatment village (intent to treat). Regressions include village × cohort fixed effects rather than individual fixed effects. Cohorts included are the following, defined in terms of current age at the time of each survey round: 0-6 months, 7-12 months, 13-18 months, 19-24 months, 25-30 months, 31-36 months, 37-42 months, 43-48 months, 49-54 months, and 55-60 months. Baseline controls not shown include the age, age², sex, and education level of the household head. Also included are dummies for gender, household size, having an improved roof, having an improved toilet, having an improved floor, having piped water, village population, and the first principal components from a PCA using information on ownership of 13 household assets at baseline. BMI is body mass index and MUAC is middle upper-arm circumference. Children with z-scores less than -6.0 or greater than 6.0 were excluded from the analysis; 59 of 1,246 height-for-age z-scores were excluded; 53 of 1,260 weight-for-age z-scores were excluded; 11 of 1,093 weight-for-height z-scores were excluded; and 14 of 1,090 BMI-for-age z-scores were excluded. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

Table A12: Robustness: multiple hypothesis testing

Treatment estimate	Outcome	Table	Column	Estimate	P-value	BKY***	BH**	Bon.*
<i>Panel A: Effects of treatment on health clinic visits of the treated in the last 12 months</i>								
Treatment (T) × 2011	0-5 years	2	2	2.30	0.01	0.02	0.02	0.06
T × 2011	60 and over	2	3	1.08	0.00	0.02	0.02	0.02
T × 2011 × fewer staff/capita	0-5 years	8	2	2.98	0.01	0.02	0.02	0.03
T × 2011 × fewer staff/capita	60 and over	8	3	0.88	0.07	0.03	0.08	0.42
T × 2011 × more staff/capita	60 and over	8	3	1.28	0.01	0.02	0.02	0.09
<i>Panel B: Effects of treatment on health clinic visits compliance in the last 12 months</i>								
T × 2011	0-5 years	3	2	0.31	0.00	0.00	0.00	0.00
T × 2011	60 and over	3	3	0.26	0.00	0.00	0.00	0.00
T × 2012	Full sample	3	1	0.04	0.03	0.04	0.04	0.08
T × 2012	60 and over	3	3	0.09	0.01	0.04	0.04	0.04
<i>Panel C: Effects of treatment on take-up of health-related products</i>								
T × 2011	Dummy - owns shoes	4	1	0.18	0.00	0.00	0.00	0.00
T × 2012	Dummy - owns shoes	4	1	0.18	0.00	0.00	0.00	0.00
T × 2012	Dummy - owns slippers	4	2	0.08	0.03	0.02	0.04	0.19
T × 2011	Insurance expenditures (thousands Tsh)	4	3	1.18	0.00	0.00	0.00	0.00
T × 2012	Insurance expenditures (thousands Tsh)	4	3	1.52	0.00	0.00	0.00	0.00
T × 2012	Dummy - participates in the CHF	4	4	0.36	0.00	0.00	0.00	0.00
T × 2011 × moderately poor	Dummy - owns shoes	9	1	0.08	0.08	0.09	0.12	0.47
T × 2011 × extremely poor	Dummy - owns shoes	9	1	0.30	0.00	0.00	0.00	0.00
T × 2012 × extremely poor	Dummy - owns shoes	9	1	0.25	0.00	0.00	0.00	0.00
T × 2012 × moderately poor	Dummy - owns shoes	9	1	0.12	0.03	0.02	0.04	0.19
T × 2011 × extremely poor	Dummy - owns slippers	9	2	0.10	0.07	0.09	0.12	0.41
T × 2012 × extremely poor	Dummy - owns slippers	9	2	0.16	0.00	0.00	0.00	0.01
T × 2011 × extremely poor	Insurance expenditures (thousands Tsh)	9	3	1.40	0.00	0.00	0.00	0.00
T × 2011 × moderately poor	Insurance expenditures (thousands Tsh)	9	3	0.87	0.00	0.00	0.00	0.00
T × 2012 × extremely poor	Insurance expenditures (thousands Tsh)	9	3	1.75	0.00	0.00	0.00	0.00
T × 2012 × moderately poor	Insurance expenditures (thousands Tsh)	9	3	1.22	0.00	0.00	0.00	0.01
T × 2012 × moderately poor	Dummy - participates in the CHF	9	4	0.33	0.00	0.00	0.00	0.00
T × 2012 × extremely poor	Dummy - participates in the CHF	9	4	0.38	0.00	0.00	0.00	0.00
<i>Panel D: Effects of treatment on illness and injury in the last month</i>								
T × 2012	Dummy - ill or injured	6	2	-0.11	0.09	0.44	0.30	0.83
T × 2012	# of days unable to do normal activities	6	4	-0.43	0.05	0.10 ⁻	0.12	0.46
T × 2012	# of days unable to do normal activities	6	5	-0.76	0.04	0.10 ⁻	0.11	0.34
T × 2012 × more staff/capita	Dummy - ill or injured	8	1	-0.06	0.07	0.44	0.30	0.64
T × 2012 × more staff/capita	# of days unable to do normal activities	8	4	-0.96	0.00	0.01	0.01	0.01
T × 2012 × more staff/capita	# of days unable to do normal activities	8	5	-1.13	0.09	0.13	0.16	0.79
T × 2012 × more staff/capita	# of days unable to do normal activities	8	6	-1.50	0.01	0.03	0.04	0.07
<i>Panel E: Effects of treatment on ability to perform ordinary activities</i>								
T × 2012	walk more than 100m	A10	5	-0.01	0.02	0.14	0.12	0.12

table continued on next page...

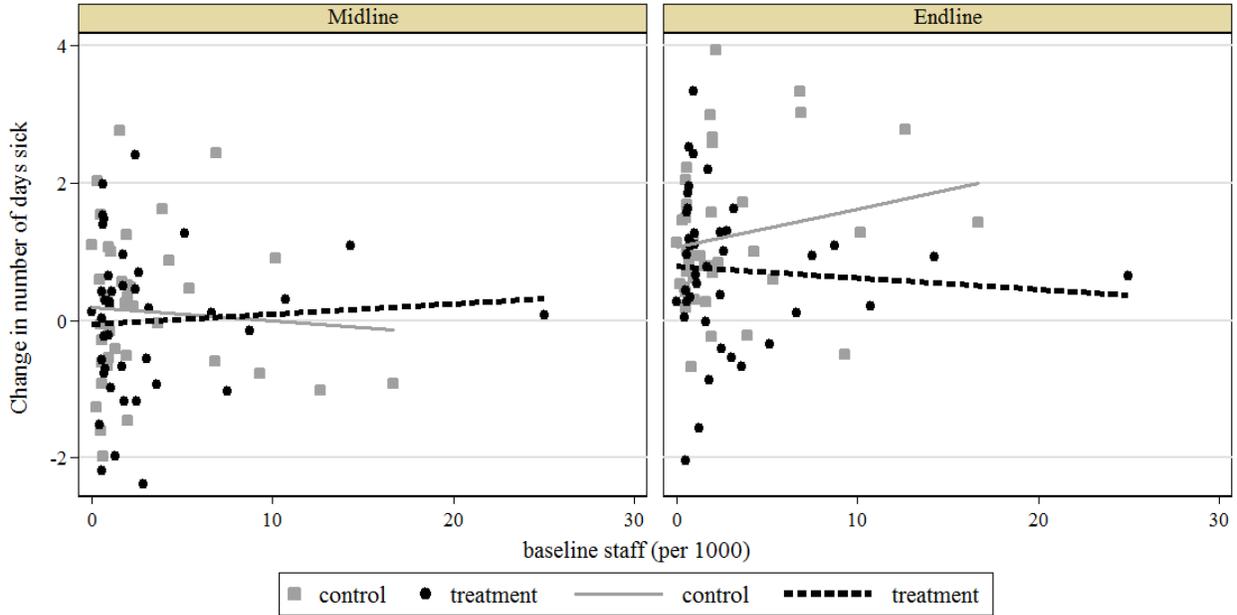
Treatment estimate	Outcome	Table	Column	Estimate	P-value	BKY***	BH**	Bon.*
<i>Panel F: Effects of treatment on type of health provider visited to address main health problem of the last month</i>								
T × 2011	None	7	1	-0.12	0.00	0.04	0.04	0.12
T × 2011	Public: Dispensary	7	2	0.17	0.00	0.01	0.01	0.01
T × 2012	Public: Dispensary	7	2	0.15	0.00	0.11	0.10 ⁻	0.10 ⁻
T × 2011 × extremely poor	None	9	1	-0.10	0.09	0.22	0.21	2.08
T × 2011 × moderately poor	None	9	1	-0.14	0.01	0.04	0.05	0.19
T × 2012 × moderately poor	None	9	1	-0.09	0.08	0.74	0.47	1.86
T × 2011 × moderately poor	Public: Dispensary	9	2	0.14	0.07	0.22	0.21	1.63
T × 2011 × extremely poor	Public: Dispensary	9	2	0.21	0.00	0.04	0.04	0.10 ⁺
T × 2012 × moderately poor	Public: Dispensary	9	2	0.13	0.07	0.74	0.47	1.72
T × 2012 × extremely poor	Public: Dispensary	9	2	0.17	0.01	0.11	0.10 ⁻	0.20
T × 2011 × moderately poor	Public: Health center	9	4	0.06	0.09	0.22	0.21	2.14
T × 2011 × moderately poor	Private: Pharmacy or chemist	9	5	-0.10	0.08	0.22	0.21	1.97
T × 2011 × moderately poor	Private: Healer, herbalist, or faith healer	9	6	0.05	0.04	0.18	0.18	1.06
T × 2011 × extremely poor	Private: Dispensary, hospital, clinic, or store	9	7	-0.04	0.05	0.18	0.18	1.10
<i>Panel G: Effects of treatment on method used to finance healthcare when addressing main health problem of the last month</i>								
T × 2012	Free treatment	5	1	-0.08	0.09	0.10 ⁺	0.16	1.09
T × 2011	Cash or asset	5	3	-0.18	0.01	0.02	0.02	0.07
T × 2012	Cash or asset	5	3	-0.18	0.01	0.01	0.01	0.07
T × 2011	Health insurance	5	4	0.16	0.00	0.02	0.02	0.03
T × 2012	Health insurance	5	4	0.28	0.00	0.00	0.00	0.00
T × 2012 × extremely poor	Loan or assistance	9	6	-0.07	0.09	0.10 ⁺	0.16	1.06
T × 2011 × moderately poor	Cash or asset	9	7	-0.21	0.00	0.02	0.02	0.06
T × 2012 × moderately poor	Cash or asset	9	7	-0.21	0.00	0.01	0.01	0.03
T × 2011 × extremely poor	Health insurance	9	8	0.16	0.02	0.03	0.05	0.24
T × 2011 × moderately poor	Health insurance	9	8	0.15	0.00	0.02	0.02	0.04
T × 2012 × moderately poor	Health insurance	9	8	0.23	0.00	0.00	0.00	0.00
T × 2012 × extremely poor	Health insurance	9	8	0.34	0.00	0.00	0.00	0.00

Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

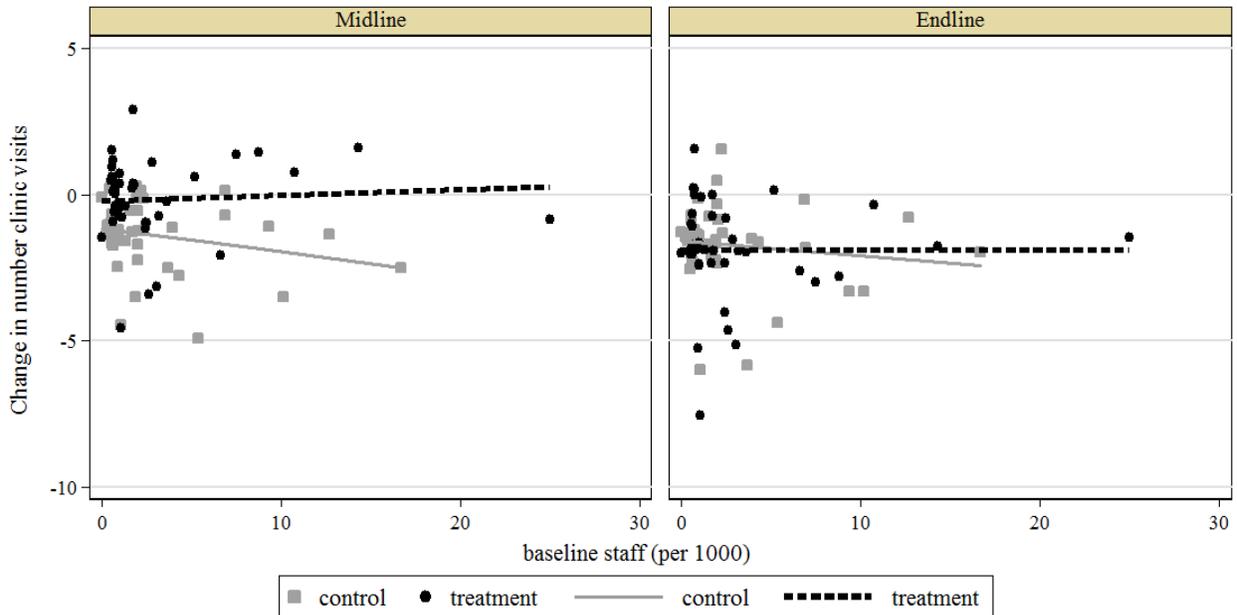
Notes: Last month refers to last four weeks. Treatment effect estimates with p-values < 0.10 displayed. Midline and endline treatment effects are abbreviated T × 2011 and T × 2012, respectively. Column refers to the column in which the estimate appears in the original table. ***BKY stands for Benjamini, Krieger, and Yekutieli q-values. **BH stands for Benjamini and Hochberg q-values. Q-values represent the smallest level at which the hypothesis is rejected. *Bon. stands for Bonferroni p-values (p-value × number of hypotheses). Of the 61 estimates with p-values < 0.10, 45, 43, and 35 estimates are significant using BKY, BH, and Bon., respectively. For q-values that round to 0.10, a plus/minus sign is provided to indicate whether the value is above/below (respectively) 0.10. BKY, BH, and Bon. are estimated separately for each period and table. The three exceptions to this rule are Table 4, Table 6, and Appendix Table A10. In Table 4, the two individual-level outcomes (owning shoes and slippers) are grouped separately from the two household-level outcomes (insurance expenditures and CHF participation). In Table 6, the two outcomes were grouped separately (one group for columns 1-3 and another group for columns 4-6; one measures the extensive margin of illness, while the other measures the intensive margin). In Appendix Table A10, the 6 specific activities of daily living are considered separately from the ordinary activities index; estimation of an index serves as an additional check on our corrections for multiple inference. Hypotheses associated with heterogeneous treatment effects, if estimated, are grouped with the hypotheses of overall treatment effects, despite being displayed in separate tables. We did not estimate q-values for the appendix tables.

Figure A1: Heterogeneous treatment impacts by health clinic staff per capita at baseline

(a) Change in the number of sick days in the last month



(b) Change in the number of clinic visits in the last year

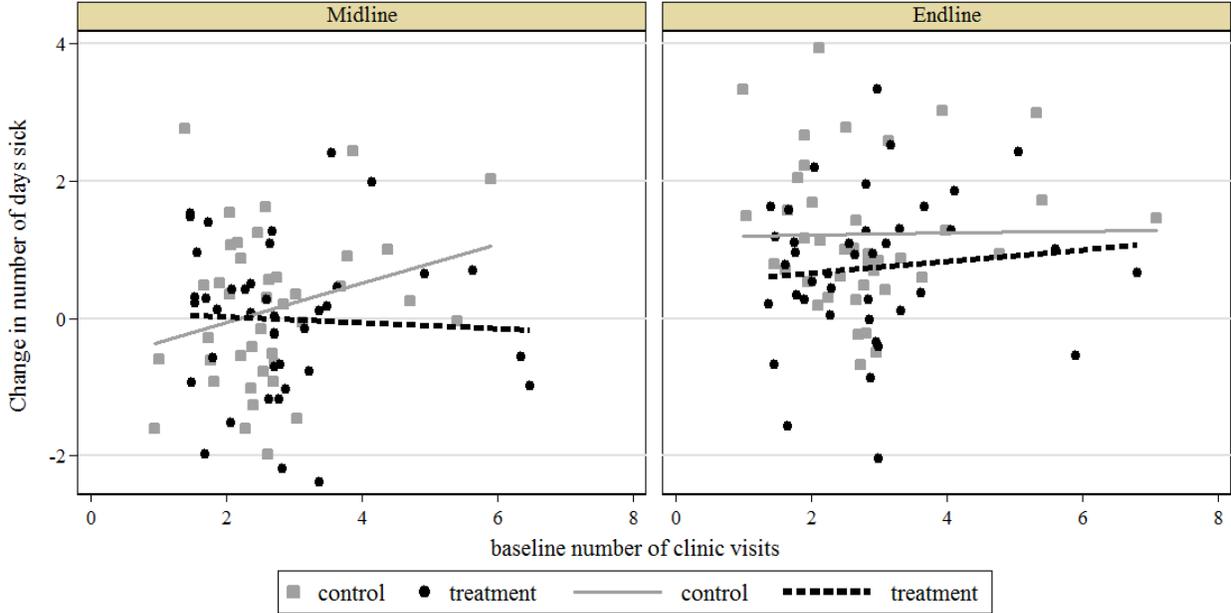


Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

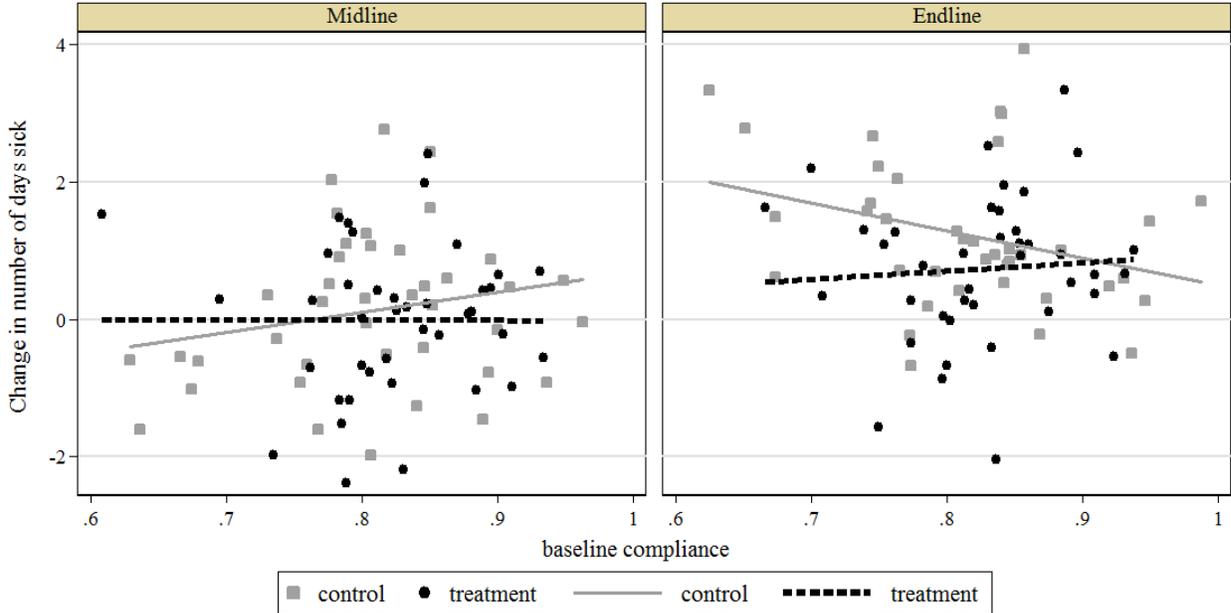
Notes: Each scatterplot contains 80 data points: 40 showing treatment village averages, and 40 showing control village averages, with separate linear fits for each set of 40. The y-axis presents the average change in the number of sick days in the last month between baseline and follow-up (positive numbers indicate increases). Since we lack midline clinic visit data for those 5-59 years, those individuals are excluded from both midline and endline plots in sub-figure (b) so that midline and endline plots are comparable.

Figure A2: Heterogeneous treatment impacts on the number of days sick in the last month

(a) By average number of clinic visits at baseline

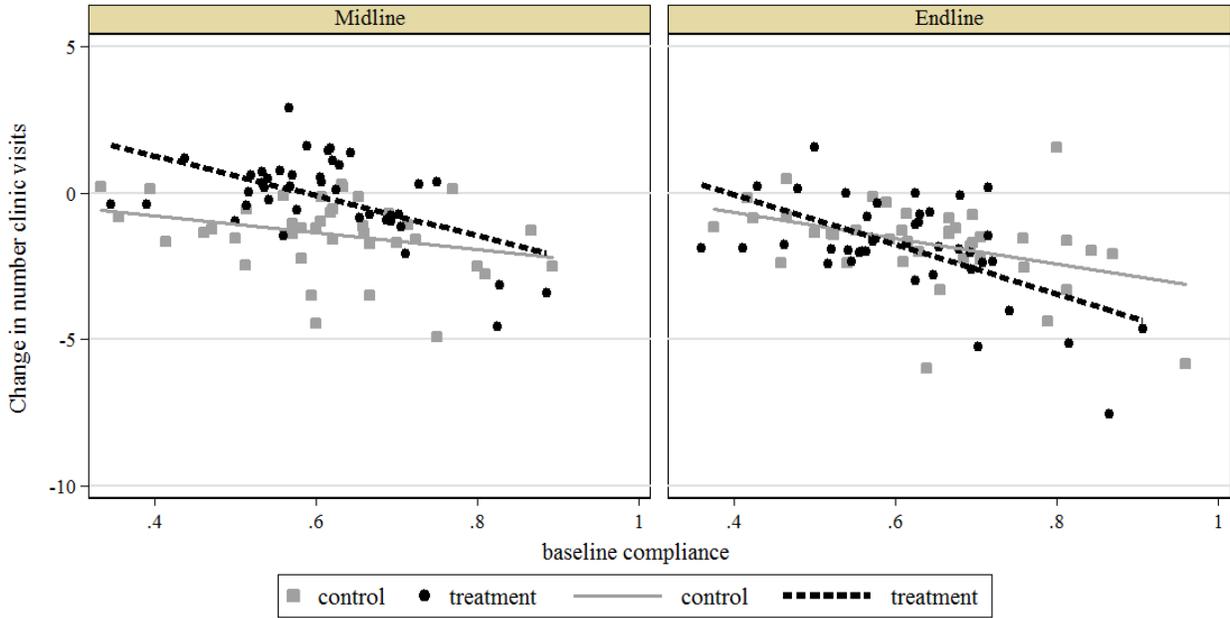


(b) By baseline rate of compliance with health clinic conditions



Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.
 Notes: Each scatterplot contains 80 data points: 40 showing treatment village averages, and 40 showing control village averages, with separate linear fits for each set of 40. The y-axis presents the average change in the number of sick days in the last month between baseline and follow-up (positive numbers indicate increases).

Figure A3: Heterogeneous impacts on number of health clinic visits by share of village in compliance with health clinic visit conditions at baseline



Source: Authors' calculations based on 2009, 2011, and 2012 household survey data.

Notes: Each scatterplot contains 80 data points: 40 showing treatment village averages, and 40 showing control village averages, with separate linear fits for each set of 40. The y-axis represents the average change in the number of clinic visits in the last month between baseline and follow-up (positive numbers indicate increases).