

Evaluating the effectiveness of a community-managed conditional cash transfer program in Tanzania

David K Evans, World Bank

Brian Holtemeyer, International Food Policy Research Institute

Katrina Kosec, International Food Policy Research Institute

Grantee Final Report
Accepted by 3ie: August 2016



Note to readers

This impact evaluation has been submitted in partial fulfilment of the requirements of grant TW1.1001 issued under Social Protection Thematic Window. 3ie is making it available to the public in this final report version. All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Katrina Kosec at K.Kosec@cgjar.org

Suggested citation: Evans, D, Holtemeyer, B and Kosec, K, 2016, *Evaluating the effectiveness of community-managed conditional cash transfer program in Tanzania*, 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie)

UKaid through the Department for International Development provided the funding for the studies carried out through the Social Protection Thematic Window.

A complete listing of all of 3ie's donors is available on the [3ie website](#).

Contents

1	Introduction	8
2	Intervention, Theory of Change, and Research Hypotheses	9
2.1	Intervention.....	9
2.2	Hypotheses and Outcomes of Interest.....	12
2.3	Theory of Change.....	14
3	Context	15
3.1	Background	15
3.2	Selection of Study Sites	16
3.3	Description of Study Sites.....	17
4	Timeline	18
5	Evaluation: Design, Methods, and Implementation	19
5.1	Ethics.....	19
5.2	Evaluation Strategy and Identification.....	19
5.3	Sample Size Determination	20
5.4	Sampling Design	21
5.5	Data Collection.....	22
5.6	Avoiding Bias and Quality Control	22
5.7	Outcome of the randomization	23
5.8	Attrition	26
6	Program: Design, Methods, and Implementation	27
6.1	Key Elements of the Program.....	27
6.2	Monitoring System.....	28
7	Results and Discussion	29
7.1	Health.....	30
7.2	Child Education and Activities	40
7.3	Household Assets	43
7.4	Expenditures and Finance	48
7.5	Community relations; data and outcomes	53
7.6	Trust and Transfers	53
8	Conclusions and Policy Implications	56
	Appendix B: Pre-analysis Plan	64
	Appendix B: Stata analysis code	68
	Appendix C: Questionnaires	70

List of Figures

1	Map of Project Areas	10
---	----------------------------	----

List of Tables

1	Timeline for implementation of CCT and accompanying impact evaluation	19
2	Baseline balance	24
3	Attrition after baseline survey	26
4	Effects of treatment on clinic visits	32
5	Effects of treatment on on take-up of health-related products	35
6	Effects of treatment on illness and injury in the last month	37
7	Effects of treatment on activities of daily living.....	38
8	Effects of treatment on anthropometrics for children aged 0-5	40
9	Effects of treatment on education	41
10	Effects of treatment on childrens' activities.....	43
11	Effects of treatment on household assets.....	45
12	Effects of treatment on livestock ownership	47
13	Effects of treatment on household savings and credit	49
14	Effects of treatment on non-food expenditures	51
15	Effects of treatment on food consumption	52
16	Effects of treatment on community trust.....	54
17	Effects of treatment on transfers over last 12 months	55

Acknowledgements

This project is the result of great effort on the part of a large number of parties. Although the authors of the report are David Evans (World Bank), Katrina Kosec (International Food Policy Research Institute, or IFPRI), and Brian Holtemeyer (IFPRI) this impact evaluation benefited at various stages from experts at the World Bank, IFPRI, TASAF, and elsewhere.

At the Tanzania Social Action Fund (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 in the evaluation, and his entire team has contributed with substantive and logistical support to the evaluation.

From the World Bank, the program and evaluation have benefited from several World Bank task team leaders, particularly Samantha de Silva, Anush Bezhanyan, and Ida Manjolo, as well as other operational staff, including Myrtle Diachok and Manuel Salazar. Berk Özler designed the survey for the TASAF II vulnerable groups evaluation, on which the household survey for this evaluation was based. Janmejay Singh consulted on the community scorecards supplementary intervention, and Julie Van Domelen consulted on the focus group questionnaires and other elements of the qualitative evaluation. Margaret Grosh and Dena Ringold provided extensive and helpful comments as peer reviewers. Arianna Legovini, who leads the World Bank's Development Impact Evaluation Initiative, provided helpful guidance.

This project is being implemented by TASAF. This survey was supported by the Japan Social Development Fund (JSDF). The impact evaluation has also received support through the Trust Fund for Environmentally & Socially Sustainable Development (TFESSD), the Spanish Impact Evaluation Fund (SIEF), and the International Initiative for Impact Evaluation (3ie).

Summary

Given the success of conditional cash transfer (CCT) programs elsewhere in the world, in January 2010 the Government of Tanzania rolled out a CCT program in three relatively poor districts: Bagamoyo, Chamwino, and Kibaha. The program was led by the TASAF. Its aim was to see if, using a model that relied heavily on communities to target beneficiaries and deliver payments, the program could improve outcomes for the poor the way centrally-run CCT programs have in other contexts. Given scarce resources, TASAF randomly selected 40 villages out of 80 eligible villages in the three study districts to be treated under the pilot program. Communities selected the most vulnerable households to participate before learning which villages were randomly selected to participate in the program. This provided a group of comparison households in the 40 untreated villages.

The program provided benefits for poor households based on the number of vulnerable children (age 0-15) and elderly (age 60+) therein. Payments were made every other month, or six times each year. While CCT payments—which averaged about US \$14.50 per month, or about 13 percent of total expenditures—were made at the household level, conditions applied at the individual level. Children aged 0 – 5 were required to visit a health clinic at least six times per year, elderly aged 60 and over were required to visit at least once per year, and no health conditions applied to other individuals. Children aged 7-15 were required to enroll in primary school and maintain an 80 percent attendance record. Locally-elected community management committees monitored compliance with these conditions and penalized participating households that did not comply by docking payments or in extreme cases removing households from the program.

A baseline survey was carried out in early 2009. Transfers began in January 2010. A midline survey was carried out in mid-2011 (18-21 months after transfers began). An endline survey was carried out in late 2012 (31-34 months after transfers began). Treatment and comparison households were broadly comparable at baseline, with few significant differences across a wide range of characteristics. In the final analysis, we compared changes over time in treatment and comparison households (a method called difference-in-differences) to adjust for small baseline differences. From these survey data, we found that the pilot program had the following major impacts.

1. After an initial surge in clinic visits among treatment households, 31-34 months into the program (at endline), participating households were attending clinics less often but were healthier: their members were sick 0.4 fewer days per month (averaging across all ages), and children age 0-5 were sick 0.8 fewer days per month.
2. Health improvements due to the CCT program are even more marked for villages with above-median clinic staff. They experienced almost an entire day per month reduction in sick days (averaging across all ages).
3. In education, the program showed clear positive impacts on whether children had ever attended school. Through qualitative data collection exercises, communities reported that the program had dramatic, positive impacts on school attendance. While these positive impacts

on absenteeism were not observed in the quantitative data, only 12 percent of children were reported to be absent during the previous week at baseline, so student absenteeism may not be a major problem. Furthermore, the program's conditions only required 80 percent attendance at school. Since attendance records and our baseline survey were collected after conditions were designed, it may have not been clear in advance that they would in many cases be non-binding. In particular, conditions were designed in part given what was considered reasonable attendance, rather than based on past attendance rates.

4. In addition, literacy rates increased significantly for girls 5-18 years old in both follow-up surveys.
5. Some of the most consistent changes observed have to do with health insurance. Treatment households were much more likely to finance medical care with insurance and much more likely to purchase insurance than were their comparison counterparts. This is important because having health insurance can substantially reduce out-of-pocket expenditures for medical care and increase the propensity to seek treatment for health problems.
6. Increases in expenditures, either on food or non-food household items, are not significantly higher for treatment households, with the exceptions of insurance and children's shoes. Households, on average, are much more likely to purchase children's shoes. This is especially true for the poorest households.
7. Treated households invested in more livestock assets. Focus groups revealed that households purchased chickens and other animals and used them to create businesses (e.g., selling eggs or chicks) or in order to have easily sellable, productive savings.
8. Because this program relies so heavily on communities — to target, to deliver transfers, and to monitor compliance with conditions — there was concern as to its impact on community cohesion. In fact, treatment households were more likely to express trust in their leaders.
9. On the whole, the community-managed CCT program led to improved outcomes in both health and education. Households used the resources to invest in livestock, in children's shoes, in insurance, and — for the poorest households — in increased savings. This suggests that the households focused on reducing risk and on improving their livelihoods rather than principally on increasing consumption. There is also evidence that the project had positive effects on community cohesion.

1 Introduction

Can a community-managed conditional cash transfer program reduce poverty in Sub-Saharan Africa? Evidence from around the world suggests that conditional cash transfers (CCTs) can effectively alleviate extreme poverty and improve a range of human capital outcomes for children (Fiszbein and Schady, 2009; Independent Evaluation Group, 2011). In recent years, evidence from Africa has shown similarly positive results (Duflo, 2000, 2003; Baird et al., 2011, 2013; Akresh et al., 2014). As the evidence on cash transfers has grown, countries around the world have raced to adopt these programs. Almost every country in Latin America now has a CCT program (Fiszbein and Schady, 2009). Garcia and Moore (2012) report that, as of 2010, at least 35 countries in Sub-Saharan Africa had implemented some sort of cash transfer program, with 14 making transfers conditional on actions taken by the recipients. However, existing CCTs have typically relied on a strong administrative role of the central government in several key aspects of program management—particularly in making payments and monitoring conditions. This program stands out as these latter two tasks were primarily led at the local level by elected community management committees (CMCs).

Giving community groups or local governments control over planning resources and investment decisions has already been shown to improve the effectiveness and efficiency of service delivery in other contexts (Chase and Woolcock, 2005; Arcand and Bassole, 2007). In some settings, however, benefits have been limited due to a problem that Mansuri and Rao (2013) refer to as a “civil society failure.” The community-driven development (CDD) approach is an innovative and untried model in the area of social protection, making it unclear whether or not it could prove successful. While such a model could potentially circumvent governance constraints to the effective operation of CCTs and social protection programs more broadly, it could potentially face a host of problems. For example, it might fail to induce compliance with conditions, generate severe leakage of funds that reduces the positive income effects of transfers, or erode communal trust and the quality of informal risk-sharing networks.

This report evaluates a new model of CCTs that relies heavily on local communities to administer many aspects of the program (e.g., targeting beneficiaries, checking compliance with conditions, and making payments) — at least to a greater degree than for past CCT programs. Lessons learned from this evaluation will thus be useful for understanding whether these aspects of operating a CCT can be effectively decentralized

Tanzania’s pilot CCT program began in January of 2010 and has continued to provide payments through the present. Payments were made every other month, or six times each year, and averaged about US \$14.50 per month (about 13 percent of total expenditures). Payments were made to beneficiaries themselves (if they were adults) or to the parent or designated caregivers of child beneficiaries. Control communities received payments starting in late 2012, shortly following the endline survey; a complete timeline of the program can be found in Section 4.

The pilot CCT program was implemented by the TASAF which was keenly interested in understanding the program’s impacts and how its design might be improved for an eventual scale-up of the program. This made a mixed methods impact evaluation critical. The principal goals of the CCT 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 operated in three districts—Bagamoyo, Chamwino, and Kibaha—where 80 eligible study villages were randomized into treatment and control groups of 40 villages each, stratified on village size and district. Random selection of villages was done after potential beneficiary households were identified in all 80 villages, to ensure comparability between vulnerable households identified in the treatment and control villages.

We evaluate the impact of the CCT program using three waves of data collected on a random subset of the beneficiary households identified in each of the 80 program villages. 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, or about 1.5 years after transfers began) and an endline survey was conducted during August–October 2012 (31–34 months, or about 2.5 years 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 midline and endline) employing focus group discussions and in-depth interviews.

Our research team prepared a pre-analysis plan (PAP) between the midline and endline surveys. While we refer to this as a PAP, it was not a PAP in the true sense. Rather, it was prepared after we had analyzed much of the midline data and written a full report using those data. As such, the typically purpose of having a PAP — to tie the researcher’s hands as far as specification and outcomes — is not served by our PAP. There is one aspect of the PAP, however, which we wrote before carrying out the analysis: a list of the heterogenous treatment effects we aimed to examine. The PAP in particular points to three main heterogenous treatment effects of interest: a) effects by exposure to severe adverse economic and climate shocks, b) effects by baseline quality of public service delivery, and c) effects by poverty level. The PAP also mentions the possibility of examining heterogenous impacts by gender and by age. We examine each of these in the course of the report; each regression examines heterogenous impacts by at most one of these.

The remainder of the report is structured as follows: Section 2 describes the intervention we evaluate, our major outcomes of interest, our hypotheses, and the theory of change. Section 3 describes the study context and addresses issues of external validity. Section 4 provides the timeline for our quantitative and qualitative data collection exercises and the roll-out of the intervention itself. Section 5 outlines the study design, datasets, identification, and measures to ensure data quality. Section 6 provides further details on the intervention and monitoring system. Section 7 presents our study findings and discusses them, taking up questions of internal and external validity. Finally, Section 8 concludes and discusses some of the policy implications of our work.

2 Intervention, Theory of Change, and Research Hypotheses

2.1 Intervention

Tanzania’s pilot CCT program began in January of 2010. Its principal goals 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 operated in three districts—Bagamoyo

(70 km from Dar es Salaam), Chamwino (500 km from Dar es Salaam), and Kibaha (35 km from Dar es Salaam)—shown in Figure 1. 80 eligible study villages were randomized into treatment and control groups of 40 villages each, stratified on village size and district. In other words, among communities of a similar size and in the same district, each community had an equal likelihood of becoming a treatment community (i.e., the potential beneficiaries identified would receive cash transfers during the evaluation phase of the project) or becoming a control community (i.e., the potential beneficiaries would not receive cash transfers during the evaluation phase of the project). The randomization methodology maximizes the likelihood that treatment and control communities are similar in unobserved characteristics as well as in measured characteristics. As random selection of villages was done after potential beneficiary households were identified in all 80 villages, this ensured comparability between vulnerable households identified in the treatment and control villages.

In each village, a community management committee (CMC) comprised of 6–14 members of the community was democratically elected and responsible for selecting beneficiaries and operating the program.¹ In the midline survey, 58 percent of households reported that a CMC member was a neighbor, and 23 percent reported that a CMC member was a blood relative. Each CMC received financial training and successfully managed at least one TASAF-supported project prior to the pilot. Immediately preceding the pilot, TASAF conducted an extensive communications and training program on the CCT at the regional, district, and village levels. CMCs were educated on how to identify and prioritize the poorest and most vulnerable households, and CMC members were then asked to carry out a survey of the poorest 50 percent of households. CMCs and community members understood

Figure 1: Map of Project Areas



Note: Adapted from the United Nations Cartographic Section map

that many fewer people than just those sampled would ultimately become beneficiaries of the program. They collected both objective data on households' poverty status and their own

¹Elections were held at a village meeting, under a closed ballot system.

subjective rating of the household's poverty level (is the household exceptionally poor or not?).² TASAF then used the data to carry out a means test and propose to the community a ranking of households within that village by poverty level. CMCs then finalized and on occasion modified the beneficiaries list under the oversight of the Village Council (VC) and with the endorsement of the Village Assembly (VA). On average, 23 percent of the villages' households were beneficiary households.

This oversight and validation helped promote community buy-in. Following beneficiary selection, CMCs in treatment villages continued to screen potential beneficiaries, communicate program conditions, transfer funds, and impose and enforce conditions. Most households were satisfied with their CMC. Across the midline and endline surveys, less than two percent of treatment households were asked for contributions related to the project. In the endline survey, only 12.5 percent of households expressed dissatisfaction with their CMC. Further, throughout the entire program, only 67 households filed complaints for receiving less in payments than they expected. As a result of such high levels of satisfaction, we have therefore not endeavored to show heterogeneous impacts of the program according to initial levels of satisfaction with one's CMC.

Treatment households received their first transfer payment in January 2010 and every 2 months thereafter. The amount of each transfer ranged from US \$12 to US \$36, depending on household size and composition. These figures were based on the food poverty line; the CCT provided US \$3 per month for orphans and vulnerable children up to 15 years of age (approximately 50 percent of the food poverty line) and US \$6 per month for elderly of least 60 years of age. In our follow-up surveys, the average reported payment was US \$14.50—about 13 percent of total expenditures over the same time period. Control group households became beneficiaries almost three years after the treatment households, in November 2012. Random selection of the control and treatment villages was done after potential beneficiary households were identified in all 80 villages, to ensure comparability between vulnerable households identified in the treatment and control villages.

While CCT payments were made at the household level, conditions applied at the individual level. Children aged 0 – 5 were required to visit a health clinic at least six times per year, elderly aged 60 and over were required to visit at least once per year, and no health conditions applied to other individuals. Children aged 7-15 were required to enroll in primary school and maintain an 80 percent attendance record. The CMC played a key role in monitoring conditions; they were responsible for collecting monitoring forms from health clinics and schools, updating records, delivering warnings when conditions were not met, making home visits to stay abreast of developments in beneficiary households, and conducting regular awareness sessions. A year and a half into the program, over 86 percent of beneficiary households reported that a member of the CMC had visited their household since the program began, and only 1.5 percent reported being asked for part of their transfer.

Monitoring of conditions began after the first payment was disbursed to beneficiaries in January

²TASAF met with local leaders to discuss who they considered “vulnerable,” and wished to target with the program. Following these discussions, TASAF provided broad guidelines to CMCs in all villages. Vulnerable children were defined as being abandoned or chronically ill, having one parent or both parents deceased, or having one or two chronically ill parents (e.g., with HIV/AIDS). Vulnerable elderly were defined as those with no caregivers, in poor health, or very poor. These guidelines helped the CMCs determine who should be interviewed as part of the census, and formed the basis for their subjective evaluations.

2010, and then was done every four months. The monitoring process was conducted by TASAF and the CMCs, with support from schools, health centers, and district staff. Monitoring forms were completed by schools and health centers, collected by the communities, and delivered to TASAF (through the district authorities) where monitoring data were entered into a computer database, and a payment list was generated.

If beneficiaries failed to comply with the conditions, a warning was issued to them by the CMCs. This, however, did not yet affect their payments. If after the next monitoring period (8 months after the first payment), beneficiaries still failed to comply with the conditions, payments were reduced by 25 percent and a second warning was sent. After two warnings were issued, beneficiaries that failed to comply were suspended indefinitely, but allowed to return to the program after review and approval by the communities and TASAF.

The CMCs played a key role in monitoring conditions, as they were responsible for collecting the monitoring forms from schools and health clinics, and conducted awareness sessions for the beneficiaries on a regular basis. They also made regular home visits to stay abreast of developments in beneficiary households in order to update the records as changes occurred in the households, and delivered warnings when conditions were not being met. As of midline, over 86 percent of beneficiary households reported that a member of the CMC had visited their household at some point since the program began in January 2010. About 93 percent of people claimed to have received their transfers from the community office, while 3.5 percent said that the CMC came to their house to deliver the payment, and the remainder received the payment in some other way. Households were included in the program for the duration of the pilot provided that they complied with the conditions. They could also leave or be asked to leave the program for the following reasons:

- If they chose to opt out, and informed the community management committee
- If the household no longer had an elderly person or a child under age 15 that was in primary school
- If household members failed to comply with conditions after a warning has been issued three consecutive times for children, and two consecutive times for elderly people
- If they moved permanently to another community where the program was not operating
- If the household representative had presented false information related to eligibility and/or committed fraud against the program.

In practice, few households were penalized for not meeting conditions. In both the midline and endline surveys, households were asked whether the last payment they received had been smaller than usual. In the midline and endline surveys, only 1.9 and 3.0 percent (respectively) of treatment households reported getting less than usual due to not meeting the conditions of the program.

2.2 Hypotheses and Outcomes of Interest

Our primary research hypothesis was that a community-managed CCT can improve a variety of individual and household welfare indicators. In particular, it will significantly increase the behaviors

that comprise the conditions of the program—health clinic attendance, school enrollment, and school attendance—in addition to improving a variety of other outcomes. Among these are:

- Individual-level outcomes, including program impacts on health-seeking behavior, on the health (including activities elderly people report being capable of performing or not and anthropometrics) and education of household members, activities performed by children, and on trust; and
- Household-level outcomes, including program impacts on investments in health, asset ownership (both household durables and livestock), savings, credit, consumption, and transfers

We measure health-seeking behavior using the number of self-reported clinic visits in the last 12 months. We lacked any administrative data on clinic visits that would have enabled us to analyze something other than self-reported data. Health is measured in four main ways: a) whether an individual was sick in the last month,³ b) for how many days an individual was sick in the last month (equal to 0 for those who were not sick), c) self-reported data by elderly individuals (aged 60+) on their ability to engage in six different activities (doing vigorous activity, walking up hill, bending over or stooping, walking more than 1 km, walking more than 100 meters, or using a bath or toilet), and d) anthropometrics for children under age five (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).

For education, we capture whether each child (aged 5 – 18 at baseline) is literate, ever attended school, is currently in school, passed the last national exam for which they sat, or missed school in the last week. For trust, we asked individuals if they trust leaders, people in general, and people in their community. Children's activities include seven activities in which a child may or may not have engaged in the last week, including fetching water, cutting wood, cleaning the toilet, cooking, caring for children, caring for the elderly, and receiving tutoring outside of school.

Investments in health are measured in four ways: a) child (ages 0 – 18) ownership of shoes, b) child ownership of slippers, c) household expenditure on formal insurance,⁴ and d) whether the household participates in the government-run health insurance program known as the Community Health Fund (CHF).

For assets, we considered the number of acres of land the household owns, ownership of nine different household durables (a mattress or bed, radio, bike, mobile phone, watch or clock, stove, iron, padded sofa, and unpadded sofa), and ownership of eight different types of livestock (indigenous cows, dairy goats, indigenous goats, local variety chickens, foreign variety chickens, sheep, pigs, and turkeys and ducks). Savings data include whether the household reports having a bank account and whether they report having other (non-bank) savings, while credit data include whether the household reports having borrowed money in the past year.

For consumption, we consider expenditures on non-food items in the last 12 months as well as the value of food consumption during the last week (both purchased and produced). We also examined the value of transfers into and out of the household—including by source for the case

³When we refer to illness in the last month, we are in all cases referring to the last four weeks.

⁴Data on expenditure on insurance is unfortunately not further disaggregated by type of insurance.

of transfers in (individuals, government, or NGOs) and by type of transfer for the case of transfers out (cash, food, or other in-kind).

In addition to the quantitative data from the household survey, we carried out qualitative analysis that provides complementary information on program impacts. Issues explored include the following: beneficiary views on program effectiveness and impact, perceptions of timeliness and amount of the transfers, reports of any irregularities, time use trade-offs for children, potential effects on intra-household transfers, empowerment effects (e.g., confidence, awareness, changes in household decision-making processes), motivational factors (i.e., besides cash, what might influence the decision of parents to send children to school, or the elderly to make regular health care visits?), issues around benefits and compliance directed to orphans, the elderly and other potentially vulnerable household members, work incentives, time demands on women, and changes in attitudes toward the education of girls and women.

2.3 Theory of Change

The community-managed CCT program is based on the following theory of change. The basic inputs—cash transfers within a framework of conditions requiring children’s school enrollment and attendance, children’s attendance at health clinics, and attendance of the elderly (age 60 and over) at health clinics—are expected to lead to the immediate outputs of increased household income (a direct result of the transfers) and increases in the behaviors on which transfers are conditioned (as these are incentivized). Project outcomes, then, would be increased consumption, increased school enrollment and attendance, and greater usage of health facilities for both the youth and for the elderly. Long-run impacts would include improvements in the well-being of children raised in these households—including better nutrition outcomes (possibly due to higher food consumption, or fewer and less severe bouts of illness), higher earnings for children raised in these households, as well as improved well-being for the elderly.

In general, outcomes could be influenced through either a behavioral effect or an income effect. In other words, beneficiaries may respond to the incentives created by the conditions, or they may simply employ their increased income to invest in “goods” that improve measures of well-being, independent of the conditions. The randomized assignment of treatment will only identify the net effect of treatment, but not the mechanism (behavior or income effect).

Impacts of the CCT operating through the channel of income effects should include a mix of immediate, medium-term, and longer-term impacts. Immediate impacts may include greater food (and non-food) consumption. Medium-term impacts may include health improvements like reductions in anemia, reductions in sick days, and more attentiveness at school. Longer-term impacts may include improvements in child anthropometrics or child income upon reaching adulthood.

Impacts of the CCT operating through the behavioral channel will be mostly immediate—including increased enrollment in school as well as increased attendance at school and clinics—and may in fact wear off quickly if the households exits the program for any reason. However, these impacts may be lasting ones in the medium and long term—with or without continued exposure to the CCT—to the extent that they create habits and expand parents’ information set about the benefits of education and visit health clinics.

Another important aspect of this program is the fact that it is community-run. While communities have been involved in some aspects of the management of CCT programs in the past, this program uniquely involves communities in a multitude of tasks—including and importantly in the areas of making payments and imposing conditions. To the extent that communities have been governance, institutions, and transparency of their management activities, we might expect more pronounced impacts of the CCT operating through either the income or behavioral channels. For example, the CCT is likely to have the largest behavioral facts where complying with conditions is least costly, and where penalties for violating conditions are highest. Complying with conditions is easier when schools and health clinics are nearby, well-staffed, and generally of high-quality. It is also easier when individuals were generally already meeting the conditions prior to treatment—indicating that the conditions are not actually binding in the first place. Some CMCs may also be relatively more likely than are others to pressure households to meet the conditions or to sanction them for a failure to meet the conditions. These “more active” CMCs may have higher intrinsic motivation to do their jobs, or they may simply operate in an environment in which record keeping is of higher quality and corruption (e.g., a parent asking or bribing a teacher not to record an absence from school) is lower. Further to this point, these high-quality, or “more active” CMCs may also be less likely to siphon off or withhold portions of payments to households (i.e. they may be less prone to corruption), leading to larger payments overall and thus greater impacts through the income channel.

3 Context

3.1 Background

The existing literature on the health impacts of cash transfers yields mixed results. Among studies from Africa, cash transfers have been shown to increase preventative health clinic visits for children in Burkina Faso (Akresh et al., 2014), improve physical and mental health in Malawi (Baird et al., 2013), raise maternal healthcare utilization for some mothers in Zambia (Handa et al., 2015), and improve anthropometric outcomes for girls—albeit not for boys—in South Africa (Duflo, 2003, 2000).⁵ In contrast, anthropometric and nutritional impacts in Latin America studies have been very mixed, with null impacts in some cases (Brazil: Morris et al. (2004), Ecuador: Paxson and Schady (2010), Nicaragua: Macours et al. (2008)) and positive impacts in others (e.g., Mexico, Colombia, and Nicaragua: Fiszbein and Schady (2009)). There are similar cases of null impacts on health outcomes in Africa now emerging (Zimbabwe: Robertson et al. (2013), Democratic Republic of Congo: Aker (2013), Kenya: Haushofer and Shapiro (2013)). And a global review of CCT programs found significant positive impacts on child anthropometry

⁵In Burkina Faso and Malawi, some recipients received CCTs and others received unconditional cash transfers (UCTs). In Burkina Faso, UCTs did not increase health clinic visits (Akresh et al., 2014). In Malawi, both improved mental health, although the benefits were lower for CCTs of high monetary value, perhaps because the transfers then make up a significant proportion of the household budget, increasing the stress associated with complying with conditions (Baird et al., 2013). Early results from a UCT program in Kenya likewise showed no impacts on health outcomes (Haushofer and Shapiro, 2013).

(Leroy et al., 2009) while another found small, insignificant impacts (Manley et al., 2013).⁶

In the education sector, there is also a wealth of evidence on the impacts of CCTs in the region. Compared to the literature on health, the literature on the education impacts of CCTs provides a more frequently positive picture of the potential of such programs to improve education indicators.

Cash transfers — both conditional and unconditional — across Africa have consistently shown positive impacts on education, mostly on access. Drawing on the results from randomized controlled trials, in Burkina Faso both conditional and unconditional cash transfers improved enrollment for boys, for older children, and for children with higher test scores at the outset of the program. However, conditional transfers were more effective for other children (girls, younger children, and those with lower test scores) (Akresh et al., 2013). A program targeting orphans and vulnerable children with unconditional cash transfers in Kenya had no impact on primary enrollment — which was already high at 88 percent — but significantly increased secondary school enrollment, despite this not being specifically targeted by the program (Ward et al., 2010). Another unconditional cash transfer program in Kenya increased access to education, and a similar program in Malawi reduced student absenteeism (Zezza et al., 2010). An unconditional child grant in Lesotho increase enrollment, particularly for adolescent boys. It did not affect grade progression (Davis et al., 2015). In Malawi, a cash transfer program that targeted adolescent girls significantly reduced dropout rates in both its conditional and unconditional forms, although dropout rates were only 43 percent as large for the unconditional group. The conditional transfers led to improvements in test scores, although the unconditional transfers did not (Baird et al., 2011). Three to four years after the transfers ended, enduring effects were found for girls who received conditional cash transfers and had dropped out of school at baseline (i.e., the program brought them back to school) (Baird et al., 2015). Trials in Zambia (unconditional) and Zimbabwe (both conditional and unconditional) also had positive, significant education impacts (Natali et al., 2015; Robertson et al., 2013). Quasi-experimental trials of unconditional cash transfers in Ghana and South Africa also show positive, significant results on children’s education (Handa et al., 2013; Edmonds, 2006).

3.2 Selection of Study Sites

The implementing agency for the CCT program was TASAF. TASAF was established in 2000, as part of the Government of Tanzania’s strategy for reducing poverty and improving livelihoods by stimulating economic activity at the community level. TASAF’s first phase of work (TASAF I) began in 2000 and involved overseeing community-run sub-projects (e.g., construction/rehabilitation of basic health-care facilities, schools and other small-scale infrastructure) which give local communities experience in managing funds, employing contractors and labor, monitoring, and reporting. TASAF I targeted the poorest and most vulnerable districts of Tanzania using a rigorous selection process. Regions were ranked using several indicators (poverty level, food insecurity, primary school gross enrollment ratio, access to safe water, access to health facilities, AIDS case rates,

⁶A few programs have examined more specialized cash transfer programs, linked specifically to maternal health investments or sexual behavior. Interventions in India and Nepal provided incentives for maternity services, with mixed results (Powell-Jackson et al., 2015; Powell-Jackson and Hanson, 2012). 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).

and road accessibility). Districts were then prioritized within the regions using an index of relative poverty and deprivation constructed using data from Tanzania's 1992 Income and Expenditure Survey. TASAF I was completed in 2005, having built a foundation for further community-driven development.

Beyond the broad support to communities under TASAF I, TASAF has implemented pilot interventions. One, the intervention discussed here, is the referred to by TASAF as the Community-based Conditional Cash Transfer project (what we are here referring to as the CCT), implemented in three district councils—Bagamoyo, Kibaha, and Chamwino. Within these three districts, all communities that had managed a TASAF I sub-project and therefore had experience in managing resources were eligible for the CCT. Other pilots, not evaluated here, were implemented in other districts. For example, Community Foundations—a partner project—was established in Kinondoni, Arusha, Morogoro, and Mwanza. In selecting the districts to implement these pilots, TASAF balanced need as well as the importance of distributing programs across areas. It is important to note—from an external validity perspective—that communities in this pilot study all had prior project management training. Results might not readily generalize to communities with no experience at all working together. That said, TASAF I only involved training for a single project that was managed by the community; as such, one might consider this to be a modest, up-front cost of later implementing a community-managed CCT program.

The pilot CCT leveraged the management capabilities of TASAF to oversee the program, and leveraged the capacities of community organizations—strengthened during the first phase of TASAF (TASAF I)—to implement it. Communities supported under TASAF I had already successfully managed sub-projects, making them relatively good candidates to operate a community-managed CCT.

3.3 Description of Study Sites

Given that the CCT included conditions on enrollment in school, attendance at school, and visits to health clinics, it is important to understand how Tanzania compares with other developing countries on health and education indicators. Here, we consider each sector in turn.

Tanzania is, in many respects, close to the Africa regional average in terms of health statistics. In 2012, the World Health Organization reported 17,318 malaria cases per 100,000 population in Tanzania versus 18,579 per 100,000 for Africa as a whole. Likewise, prevalence of HIV is at 3,082 (per 100,000) versus 2,774 for the region. Life expectancy at birth is 61 years versus 58 for Africa as a whole. The distribution of years of life lost across communicable diseases, non-communicable diseases, and injuries is very similar. Yet on some measures, Tanzania diverges significantly from the rest of the region. Its under-five mortality rate (54 per 1,000 live births) is just over half that of the region (95). Its maternal mortality ratio is almost twenty percent lower than the region as a whole. The health workforce, however, is weaker in Tanzania, with just 0.1 doctors per 10,000 population (versus 2.6 for Africa on average) and 2.4 nurses and midwives (versus 12.0 for Africa) ([World Health Organization, 2014](#)).

Crude measures of healthcare utilization in Tanzania—an area where cash transfers conditioned on health might be expected to have a large impact—suggest significant room for improve-

ment. Contraceptive prevalence among women aged 15–49 is only 34 percent, only 43 percent of pregnant women make a full set of antenatal care visits, and about half of births are attended by skilled birth attendants ([World Health Organization, 2014](#)). Despite the limited health workforce, recent evidence disentangles the effect of using formal public health facilities from self-selection to demonstrate significant improvements in health outcomes for children who take advantage of these facilities ([Adhvaryu and Nyshadham, 2015](#)). After user fees were introduced to facilities in the early 1990s, the Tanzanian government introduced a health insurance program called the Community Health Fund (CHF). This program is a voluntary, district-level prepayment scheme. Members pay a fixed annual fee of between 5,000 and 10,000 Tanzanian shillings (between \$3 and \$6 US⁷), depending on the region, but then their entire family is exempt from any co-payments for visits to primary healthcare facilities ([Marriott, 2011](#)).⁸ Upon introduction of the CHF, children and maternal health services were already exempt from co-payments according to official government policy ([Babbel, 2012](#)). Ten years after the introduction of the CHF, the program had an average enrollment rate of only ten percent. At least two of the reasons cited for lack of participation were inability to pay or to see the rationale to insure ([Kamuzora and Gilson, 2007](#)). Insofar as liquidity has been a binding constraint, a cash transfer program might be expected to significantly impact participation.

In the education sector, Tanzania has made great strides. Relative to the rest of the region, it performs well on certain access variables. Median primary school completion rate for the region (in 2013) was 71%, whereas Tanzania achieved a rate of 76%. At the secondary level, Tanzania is almost exactly par for the region, with a 39% low secondary completion rate, as compared to 40% for the region ([World Bank, 2016](#)). (Note that many countries have missing data, and this only includes countries for which data are available.) Access has expanded dramatically in Tanzania over the last decade, as has been the case in many parts of the continent. Primary school enrollment grew from 4.8 million in 2001 to 8.4 million in 2010 ([World Bank, 2014](#)).

However, quality of education is still a challenge. Recent reports from across East Africa have demonstrated that fewer than one third of children in third grade possess even basic literacy or numeracy skills. For seventh-grade children, one in five do not have the literacy and numeracy competencies for Grade 2. Tanzanian children has pass rates higher than Ugandan children but lower than Kenyan children ([Uwezo, 2014](#)). These results are unsurprising, given the challenges in service delivery: Recent research suggests that only 42% of teachers pass a minimum competency test, and absentee levels are extremely high ([World Bank, 2013](#)). These service delivery challenges are demonstrated in many countries. Kenya had an almost identical proportion of teachers pass a minimum competency test, with lower absenteeism from school but higher absenteeism from the classroom.

It is also important to understand how the study population compares with Tanzania as a whole. The objective of the project was to benefit relatively poor households, and comparing baseline data from households in the study to national data around the same time demonstrates that this was the case. For example, just over 80% of Tanzanian households were without electricity, whereas

⁷In 2009 the exchange rate ranged from 1,280 to 1,467 per U.S. dollar ([Bank of Tanzania, 2015](#)).

⁸Up to 7 family members are exempt from co-payments—though receipt of medications/ tests incurs fees.

nearly 100% of study households were. Study households were much more likely to have a mud floor (just over 60% for the country versus nearly 100% for the study population) and were much more likely to be headed by women.

4 Timeline

Table 1 below presents the combined chronology of both the program and the evaluation.

Table 1: Timeline for implementation of CCT and accompanying impact evaluation

Timing	Activity
November 2007 - September 2008	Program Design (completion of Operational Manual, set up of MIS, preparation of guidelines, forms, and materials for training activities)
September - November 2008	Sensitization at regional, district, ward, and community levels
October - November 2008	Targeting activities (field data collection, data entry, and community validation of beneficiaries)
October - November 2008	Training of district officers and community management committees on the targeting process
January - May 2009	Baseline survey
September - October 2009	Enrollment of beneficiaries
January 2010	First payments made to beneficiary households
November 2010 - February 2011	Community Scorecard Exercise
July - September 2011	Midline survey & first round of focus group interviews
August - October 2012	Endline survey
July - August 2013	Second round of qualitative data collection, including in-depth and focus group interviews

5 Evaluation: Design, Methods, and Implementation

5.1 Ethics

All research work carried out at IFPRI must be closely scrutinized by members of the IFPRI Institutional Review Board (IRB) to make sure that study methods/protocols do not contravene set standards of ethics to protect human subjects. Prior to the actual implementation of the quantitative and qualitative data collection in which IFPRI was involved (the 2012 endline survey and the 2013 qualitative data collection exercise), an IRB application and copies of survey instruments were submitted for IRB approval. At the outset of the study, the World Bank explored whether there were a national body responsible for evaluating ethical research and were informed that this was not the case. However, the World Bank team and the Government of Tanzania collaborated to ensure that all surveys included informed consent, that respondents were informed that they

could refuse to answer any questions, and data with identifying information were not available to government bodies.

5.2 Evaluation Strategy and Identification

We evaluate the impact of the CCT program using three waves of data collected on a random subset of the vulnerable households identified in each of the 80 program villages. Data were collected only for beneficiary households or—in the case of control villages—would-be beneficiary households that only did not receive transfers because their village was not randomly assigned to treatment. 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, or about 1.5 years after transfers began) and an endline survey was conducted during August–October 2012 (31–34 months, or about 2.5 years 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 midline and endline) employing focus group discussions and in-depth interviews.

We carried out a baseline and two follow-up surveys, in 2011 and in 2012, to capture both short-term (1.5 years) and medium-term (2.5 years) impacts of this CCT program. As assignment to treatment was random, we can recover causal estimates of the impacts of the CCT program by estimating the following empirical specification:

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

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 $2012_t = 1$ at the time of the endline (August - October 2012) and zero otherwise. When we consider a household-level outcome, i instead indexes households.

Ultimately, some households initially identified for treatment did not receive treatment, and some few households initially not intended for treatment did receive it. Specifically, in treatment villages, 9.0% of assigned households did not receive treatment—likely due to last minute changes in community prioritization or household refusal. In control villages 0.6% of households did receive treatment—likely due to their close proximity to a treatment village. Our standard specification, specified above, is an intent-to-treat estimate. It is more conservative than actual impact of treatment. We also estimate specifications in which we estimate the impact of actually receiving treatment, in which case we use whether or not someone was assigned to treatment as an instrumental variable to the endogenous variable of actually receiving treatment. This instrumental variable satisfies both conditions required for legitimate instrumentation: First, it is highly correlated with actual treatment, as only a small percentage of assigned households failed to receive treatment. Second, it is uncorrelated with other factors that might directly affect outcomes, as assignment was random.

5.3 Sample Size Determination

The number of communities was determined by political realities: 80 communities within the three pilot districts were eligible for the program. The number of households to be interviewed per community was determined through power calculations. With a total of 80 participating communities (40 treatment and 40 control) and an effect size of 0.20, we expected to need to interview 20 households per community in order to achieve 80% power. This assumed 95% confidence levels for statistical significance and an intra-cluster correlation of 0.05. Evaluations of conditional cash transfer programs elsewhere had found effects of this size. For effects of this magnitude on health and education outcomes in a Nicaraguan CCT ([Rawlings and Rubio, 2005](#), Table 6). For Mexico's PROGRESA (subsequently Oportunidades and then Prospera) program, see effect sizes on child height in [Behrman et al. \(2000\)](#). For effect sizes on longer-term schooling outcomes, see [Behrman et al. \(2005\)](#). We used these calculations to inform the baseline and midline surveys, seeking to interview 25 households (on average) per village, thus permitting up to 20% attrition within villages and still maintaining 80% power.

5.4 Sampling Design

The districts to be included in the pilot were selected by TASAF. Households to be included in the pilot were determined as follows. Targeting was done by community management committees (CMCs) under the oversight of the Village Council (VC), the local governing body, and with the endorsement of the Village Assembly (VA), which consists of all adults who live in the village. The CMC was democratically elected by potential beneficiaries and endorsed by the VA. Targeting was done using screening forms designed to identify vulnerable children and elderly people based on the following criteria, which were defined by the communities themselves. Vulnerable children were defined as follows:

- One parent or both parents deceased
- Abandoned children
- Having one or two chronically sick parents (for example, human immunodeficiency virus [HIV]/AIDS)
- Chronically sick children, despite having two parents alive. Vulnerable elderly were defined as follows:
 - Elderly with no caregivers
 - Poor health
 - Very poor

The CMC used these poverty indicators to identify the poorest (approximately) half of households in the community. Next, the CMC—under the supervision of local government authority (LGA) facilitators and the guidance of the VC—collected data from the identified households using

a special screening form for first verification by proxy means test. LGA facilitators then verified the accuracy of collected data. With these data, TASAF performed proxy means testing on a sample basis to ensure that targeted beneficiaries qualified.

Households were divided into three groups: eligible, ambiguous, and rejected. Validation of the list of eligible households was done by the Village Assembly, allowing for community validation. Priority ranking of households was conducted in the event that the number of beneficiaries exceeded available resources, along the following criteria:

- First priority: Households with a child as head of the household
- Second priority: Households with an elderly person as head of the household
- Third priority: Households with only elderly persons

The final list of households was then endorsed by the VA. Within this list of households in treatment and control villages, households were randomly selected for data collection. Households were sampled from the universe of communities participating in the pilot.

Treatment was assigned randomly at the village level. Individual-level treatment was discussed with the Government but within-village randomization was determined to be too politically sensitive. The contracted data collection firm, Economic Development Initiatives (EDI), carried out a simple randomization using computer software; the randomization was not carried out publicly. The comparison group was aware of the program.

5.5 Data Collection

Primary data were collected at baseline (January through May 2009), midline (July through September 2011), and endline (August through October 2012). The principal data collection instrument was a household survey, although brief instruments were administered to community leaders. The initial instrument was adapted from the baseline survey for the TASAF II vulnerable groups impact evaluation. Then, the impact evaluation team and the Tanzanian government reviewed the survey item by item to ensure that it covered all needed areas of the impact evaluation.

EDI was contracted to collect the data in all three rounds, in both treatment and comparison groups. In each case, enumerators and supervisors underwent an extensive training.

Household surveys were gathered at the home of the interviewee. All data were gathered using tablet personal computers, as these offered the advantage of automated data entry and built-in data checks (e.g., disallowing nonsensical numerical entries).

5.6 Avoiding Bias and Quality Control

There are a range of potential biases that can affect data collection and impact evaluation work. These can include, among others, social desirability bias, Hawthorne effects, John Henry effects. To avoid social desirability bias, no representatives from TASAF were present during the actual interviews. Furthermore, rather than asking households specifically about, for example, consumption of temptation goods (such as alcohol or tobacco) all at once, households were rather asked

about a long list of consumption items, without any one being singled out for special attention. Furthermore, evidence from elsewhere on misreporting (in order to overestimate poverty or to reduce social embarrassment) is especially a problem in surveys carried out away from the household (see Martinelli & Parker, “Deception and Misreporting in a Social Program,” *Journal of the European Economic Association*, 2009). All three rounds of data collection were carried out in the household and therefore those problems are expected to be minimal. A number of poverty measures, such as household improvements, are directly observed by the interviewers. Other measures, such as school attendance, are structured very specifically to avoid psychological biases that may incline people toward reporting positive behaviors “on average.”

At the same time, there is of course some concern that households in villages assigned to CCT treatment might change their behavior not due to the CCT, but merely due to being studied (Hawthorne effects). If Hawthorne effects lead people to exaggerate how often they attend health clinics or send their children to school (which are CCT conditionality requirements), this should be true across both treatment and comparison groups and so should not lead to bias across groups. If Hawthorne effects lead people in treatment villages to actually attend school or go to health clinics, this study overcomes that by the fact that we study people over a relatively long period of time and have both a midline survey (after 1.5 years of transfers) and an endline survey (after 2.5 years of transfers). Hawthorne effects are generally short-lived. It is unlikely that people would engage in behavior modification for a full 2.5 years, constantly keeping in mind that they are treatment households being studied. Our survey is a rather concise instrument, and it will only be carried out twice, at long-spread-apart intervals, following assignment to treatment. This is not the typical timespan during which Hawthorne effects have been identified in previous work.

Alternatively, John Henry effects could be a concern if households in control villages feel that they have something to prove about not needing CCTs in order to attend health clinics and schools, or in order to prove their “worthiness” for CCTs in the future. However, we consider this unlikely; if anything, we would be more worried that control households want to signal their need for CCTs (i.e., show that they have poor outcomes without CCTs) in order to attract the attention of the Government of Tanzania (which was tentatively planning a scale-up of the program). Nonetheless, like Hawthorne effects, we can expect that John Henry effects would be relatively short-lived. If this is the case, then they should not affect the results of surveys carried out 1.5 (our completed midline survey) and 2.5 years (our proposed endline survey that 3ie would be funding) after assignment to treatment.

While data entry was not a part of this study, due to electronic data collection, in the coding of the data we used a dual analysis system, in which at least one principal investigator would re-examine the work from any more junior research analyst.

5.7 Outcome of the randomization

Despite randomization of villages into treatment and control, it is possible that some observable characteristics of treatment villages were different than those of control villages before the intervention. If this were the case, one would worry that the control villages are not a reasonable counterfactual to how the treatment villages would have been without treatment. A comparison

of baseline sample means in treatment and control villages reveals balance on the vast majority of our outcomes (Table 2). Across 98 outcomes, one, eleven, and eight outcomes are significant differences at the 1, 5, 10 percent levels, respectively. However, despite the fact that there is some imbalance, the size of the differences is often small, and we do not see evidence of systematic differences across groups. Furthermore, we use individual fixed effects (or household fixed effects for outcomes that vary at the household level) to account for any baseline imbalances.

Table 2: Baseline balance

Outcome	Treatment (T)		Control (C)		Difference (T-C)	
	Mean	N	Mean	N	Mean	S.E.
Clinic Visits (Table 4)						
# health clinic visits in past year (Full Sample)	2.83	3462	2.77	3456	0.06	(0.26)
# health clinic visits in past year (0-5 years old)	8.21	309	8.33	312	-0.12	(0.70)
# health clinic visits in past year (60 and over)	2.91	1049	2.67	1160	0.24	(0.35)
Health-related products (Table 5)						
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)
Health (Table 6)						
Dummy - ill or injured in last month (Full sample)	0.29	3462	0.26	3456	0.02	(0.02)
Dummy - ill or injured in last month (0-5 years old)	0.31	309	0.25	312	0.06*	(0.04)
Dummy - ill or injured in last month (60 and over)	0.39	1049	0.38	1160	0.01	(0.03)
# of days unable to do normal activities (Full sample)	1.68	3462	1.59	3455	0.08	(0.14)
# of days unable to do normal activities (0-5 years old)	1.31	309	0.80	312	0.51**	(0.21)
# of days unable to do normal activities (60 and over)	2.78	1049	2.79	1159	-0.02	(0.31)
Specific activities of daily living, 60 and over (Table 7)						
Dummy - Do vigorous activity	0.61	2216	0.61	2278	0.00	(0.03)
Dummy - Walk uphill	0.87	2216	0.86	2278	0.01	(0.02)
Dummy - Bend over or stoop	0.98	2216	0.98	2278	0.01	(0.00)
Dummy - Walk more than 1km	0.93	2216	0.92	2278	0.01	(0.01)
Dummy - Walk more than 100m	0.98	2216	0.97	2278	0.01	(0.00)
Dummy - Use bath or toilet	0.99	2216	0.98	2278	0.01	(0.00)
Dummy - Ordinary activities index	5.36	2216	5.32	2278	0.04	(0.06)
Anthropometrics (Table 8)						
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)	156	230	156	232	-0.15	(1.42)
Education (Table 9)						
Dummy - Literate	0.56	1162	0.57	1071	-0.01	(0.04)
Dummy - Ever attended school	0.75	1162	0.78	1071	-0.03	(0.03)
Dummy - Currently in school	0.87	870	0.89	834	-0.02	(0.02)
Dummy - Passed last national exam	0.79	406	0.86	377	-0.07***	(0.03)
Dummy - Missed school	0.08	1159	0.09	1069	-0.01	(0.02)
Child activities (Table 10)						
Dummy - Fetch water	0.56	1204	0.59	1129	-0.03	(0.03)
Dummy - Cut wood	0.29	1204	0.27	1129	0.03	(0.03)
Dummy - Clean toilet	0.13	1204	0.13	1129	0.00	(0.03)
Dummy - Cook	0.25	1204	0.27	1129	-0.02	(0.02)
Dummy - Care for children	0.16	1204	0.16	1128	0.00	(0.03)
Dummy - Care for elderly	0.26	1204	0.31	1129	-0.05	(0.04)
Dummy - Receive tutoring	0.02	1203	0.03	1129	-0.01	(0.01)
Household assets (Table 11)						
Number of acres of land	4.44	877	3.64	877	0.80*	(0.40)
Dummy - mattress or bed	0.69	881	0.73	879	-0.04	(0.06)
Dummy - Radio	0.31	881	0.34	879	-0.03	(0.04)
Dummy - Bike	0.20	881	0.18	879	0.03	(0.03)
Dummy - Mobile phone	0.09	881	0.12	879	-0.03	(0.02)
Dummy - watch or clock	0.09	881	0.11	879	-0.02	(0.02)
Dummy - Stove	0.07	881	0.12	879	-0.05**	(0.03)
Dummy - Iron	0.04	881	0.07	879	-0.03**	(0.01)
Dummy - Sponged sofa	0.01	881	0.03	879	-0.02*	(0.01)
Dummy - Non-sponged sofa	0.01	881	0.02	879	-0.01**	(0.01)
Household livestock (Table 12)						
Dummy - Indigenous cows	0.11	880	0.05	879	0.06	(0.06)
Dummy - Dairy goats	0.00	880	0.00	879	0.00	(0.00)
Dummy - Indigenous goats	0.26	880	0.24	879	0.02	(0.11)
Dummy - Local variety chickens	2.34	879	2.40	879	-0.06	(0.29)

(continued on next page)

(Baseline balance continued...)

Outcome	Treatment (T)		Control (C)		Difference (T-C)	
	Mean	N	Mean	N	Mean	S.E.
Household livestock (continued) (Table 12)						
Dummy - Foreign variety chickens	0.03	880	0.18	879	-0.15	(0.17)
Dummy - Sheep	0.03	880	0.01	879	0.02	(0.02)
Dummy - Pigs	0.02	880	0.02	879	0.00	(0.01)
Dummy - Turkeys and ducks	0.20	880	0.08	879	0.12**	(0.05)
Household Savings and credit (Table 13)						
Dummy - Has a bank account	0.02	880	0.02	879	0.00	(0.01)
Dummy - Has other savings	0.01	879	0.02	878	-0.01	(0.01)
Dummy - Borrowed in past year	0.19	880	0.18	879	0.01	(0.02)
Non-food expenditures (TSH) (Table 14)						
Tobacco products	5766	879	4764	877	1002	(1218)
Children's clothing	5595	881	7185	879	-1590	(1251)
Clothing/footwear for men	5863	880	6079	877	-217	(999)
Clothing/footwear for women	8089	881	9459	876	-1370	(1474)
Other personal effects	571	880	1058	877	-487**	(210)
Weddings/funerals/dowries	3421	881	4448	878	-1027	(632)
Medical care - services	8056	880	10771	878	-2715	(1635)
Medication	4138	881	5832	878	-1693**	(846)
Boarding school costs	3458	880	5649	877	-2190	(2113)
Food expenditures (TSH) (Table 15)						
Maize (flour/super/sembe) - purchased	3542	881	3657	879	-115	(616)
Maize (flour) super/sembe - produced	182	881	250	879	-68.04	(110)
Maize(flour) dona - purchased	707	880	634	878	73.37	(171)
Maize(flour) dona - produced	421	881	321	879	99.59	(142)
Other flour (millet/cassava/sorghum/barley) - purchased	174	879	116	877	57.94	(59.11)
Other flour (millet/cassava/sorghum/barley) - produced	357	881	491	879	-134	(148)
Husked rice- purchased	170	881	357	879	-187**	(86.70)
Husked rice- produced	53.13	881	69.68	879	-16.55	(44.47)
Dried beans - purchased	471	881	631	879	-161*	(92.65)
Sugar - purchased	552	881	608	879	-56.59	(106)
Community trust (Table 16)						
Dummy - leaders can be trusted	0.81	878	0.80	873	0.01	(0.03)
Dummy - most people can be trusted	0.26	875	0.23	874	0.03	(0.03)
Dummy - community people can be trusted	0.59	876	0.53	873	0.06*	(0.03)
Transfers into household (TSH) (Table 17, Panel A)						
Total	18314	883	25401	881	-7087**	(3427)
Individuals	16846	880	23616	879	-6770*	(3479)
Government or TASAF	833	880	463	879	369	(225)
NGOs	698	880	1187	878	-489	(449)
Transfers out of household (TSH) (Table 17, Panel B)						
Total	978	879	1223	877	-246	(374)
Cash	303	879	327	878	-24.83	(135)
Food	485	880	307	875	178	(115)
Other in-kind	72.81	879	204	877	-132	(80.72)
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)
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)

Note: From authors' calculations based on baseline (2009) household survey data. Treatment indicates assignment to treatment. Standard errors are clustered at the village level, and appear below the coefficient in parentheses. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$

5.8 Attrition

Between baseline and midline, 8.6 percent of households attrited from the sample, and between baseline and endline, 13.2 percent attrited. This attrition poses two potential problems. First, if attrition is correlated with treatment status, our findings at the two follow-up surveys could merely reflect the differential sample composition between the treatment and control groups rather than the causal effects of treatment. This would indicate that attrition had compromised the internal validity of the results. Second, if attrition is greater for particular types of households—even if it does not vary with treatment status—this would compromise the external validity of the results. Our estimates of the effects of treatment would only be representative of the types of households that did not attrit from the sample, rather than the full array of poor households represented in our initial sample.

Table 3: Attrition after baseline survey

	Household remains in ... survey			
	Midline		Endline	
	(1)	(2)	(3)	(4)
	b/se	b/se	b/se	b/se
Treatment village	0.002 (0.016)	-0.119 (0.178)	0.023 (0.020)	0.174 (0.251)
Head age		0.003 (0.004)		0.012** (0.006)
Head age × Treatment		0.003 (0.006)		-0.004 (0.008)
Head age ²		-0.000 (0.000)		-0.000** (0.000)
Head age ² × Treatment		-0.000 (0.000)		0.000 (0.000)
Head has some education		-0.001 (0.019)		-0.010 (0.024)
Head has some education × Treatment		0.032 (0.029)		0.004 (0.034)
Asset index		0.012*** (0.004)		0.003 (0.008)
Asset index × Treatment		-0.010 (0.007)		0.001 (0.011)
Constant	0.913*** (0.013)	0.886*** (0.106)	0.857*** (0.015)	0.540*** (0.184)
Observations	1764	1757	1764	1757

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Treatment indicates assignment to treatment. 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 $p < 0.10$.

Fortunately, neither of these two problems materializes in either of the two follow-up surveys, as shown in Table 3. Columns (1) and (3) regress a dummy for attrition on our treatment dummy,

for midline and endline, respectively. Columns (2) and (4) regress a dummy for attrition on our treatment dummy, an array of household controls (head age, head age squared, head education level, and household asset ownership), and the interactions of these controls with our treatment dummy—once again for midline and endline, respectively. In no case does the treatment dummy significantly predict attrition. Further, in no case is the interaction between a household characteristic and treatment statistically significant. At midline, we see that richer households are slightly more likely to attrit from the sample, but this effect disappears by endline. At endline, we see that older household heads were more likely to attrit from the sample—though this effect was not present at midline. Given the large number of coefficients we test, it is not surprising that some are significant; indeed, by pure chance we might expect that to be the case. Overall, we conclude that attrition is not a problem affecting either the external or internal validity of our results.

6 Program: Design, Methods, and Implementation

6.1 Key Elements of the Program

The selection of treatment and control households followed the following process:

Phase 1: Selection of program communities. In this phase, the team compiled information for all communities in the three program districts on their population, the existence of the infrastructure necessary to accommodate the increase in demand for community education and health services that a CCT would induce, and the experience and quality of CMCs. This information was necessary to both stratify the sample and ensure that program communities were suited to the requirements of the CCT, including provision of services and enforcement of program conditions.

Phase 2: Identification of eligible households. In this phase, the potential beneficiaries in all program communities (not yet divided into control and treatment communities) were identified. CMCs and Village Councils prepared ranked lists of households based on the criteria for vulnerable households that had been previously determined in discussions with TASAF communities. These lists informed the selection of recipient households in treatment communities and of households for data collection in control communities. Expectations of residents in all 80 program villages were managed by providing clear communication from the start that not all communities could participate. (Ultimately, additional resources were secured so that transfers could be rolled out to control communities immediately after the endline survey concluded, in November 2012.)

Phase 3: Selection of the treatment and control communities. Once eligible households were identified in all 80 program communities, 40 treatment communities were selected at random. Random selection was stratified on known community characteristics (such as sub-district and community size) to ensure comparability between treatment and comparison communities.

Phase 4: Selection of the treatment and control households. The design team used the total share of the eligible population across all selected communities to ensure proper coverage

among all treatment communities. CMCs received a cap on how many households in the community could participate in the program based on a combination of community population and poverty map projections.

Phase 5: Data collection. There were many more program beneficiaries (in treatment communities) and potential beneficiaries (comparable households in control communities) than could feasibly be interviewed. Once all communities were assigned into the treatment or comparison groups, power calculations identified the need to interview an average of 25 households per community. In cases where participating households (i.e., households that would receive treatment, whether in a treatment or control community) did not exceed that number, the team interviewed the full sample of target households in that community. In communities with more than 25 participating households, the team collected data on a random sample of 25 households.

6.2 Monitoring System

Monitoring activities for this pilot fell into two major categories: routine monitoring and a community score card exercise.

1. Routine monitoring and reporting activities: These were carried out as part of implementation by TASAF and local government authorities, with input from communities, to ensure that activities were being carried out as planned, proper targeting had taken place, and funds were properly disbursed. TASAF submits quarterly financial management reports, and conducts semi-annual audits of community accounts. TASAF is subject to independent financial audits led by Tanzania's Auditor General, and also undertakes systematic process and technical audits (all of which have been highly satisfactory to date). Information provided by the community management committees on monitoring of conditions was randomly cross-checked against submissions from the schools and health facilities.
2. Community Score Cards: A module on Community Score Cards (CSCs) was used as part of the intervention itself to enhance the accountability and process monitoring of the CCT roll out. CSCs are simple community monitoring tools that blend different participatory monitoring approaches and social accountability techniques (such as social audits and citizen report card surveys). They have proven to be powerful instruments to exact accountability and promote transparency in rural contexts. The CSC process consists of four elements:
 - a Input tracking — in which a mini social audit is undertaken at the community level that attempts to match project/program inputs with actual outputs and disbursement. In the context of the CCT pilot it means tracking disbursements and timing of CCTs to stated beneficiaries and cross-checking targeting efficiency. For the schools and health centers themselves, it tracks key infrastructure and materials that are available (e.g., classrooms, medicines, and medical equipment);
 - b Community performance scorecard — in which different focus groups (e.g., CCT beneficiaries, non-beneficiaries, youth, elders, men, women) in each community rate the

performance of different elements of a program (in this case the CMC management, CCT system, or the school and health facilities participating) on different performance criteria (this could include criteria such as transparency, fairness, timeliness, or adequacy), as well as the services being provided (e.g., whether teachers are present, as well as health personnel, supplies, medicines);

- c Self-evaluation scorecard — the community management committee that is administering the CCT and the schools and health centers participating in the program themselves give a self-assessment of how they see the system performing (these could be similar to the criteria above, but normally one finds that providers rate themselves differently compared to beneficiaries); and
- d The interface meeting — providers (CMC, health staff, and school teachers) and the community are finally brought together to share their results, discuss the findings, and jointly plan on how to make the process work better. This action plan can then feed back to TASAF management and ideally would help modify the operation of the pilot in subsequent rounds.

Program protocols were generally preserved, without deviation. As implementers had received prior training during phase I of TASAF, they were already familiar with handling funds and communicating information to TASAF. We know of two innovations made by implementers in the course of the evaluation: our qualitative work suggested that at least in one village, health clinic workers traveled to the location in which transfers were handed out to encourage beneficiaries to sign up for the CHF (which would give free clinic visits to up to seven members of their family). Further, our qualitative work suggested that many teachers and CMCs would travel to beneficiaries' households to persuade and pressure them into compliance with the CCT program conditions. Such innovations may have increased the effectiveness of the intervention.

Study participants were aware that they were part of an experiment. Before randomization of communities into treatment and control, potential beneficiaries in all communities were made aware—via community meetings—that they had the possibility of receiving a CCT starting in early 2010, but that it was possible that they would not receive the program until late 2012. Further, during our baseline survey, sample households learned that they were being studied.

7 Results and Discussion

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 education outcomes, we examine impacts by gender. Parents often make differential education investments in their children according to gender, and it is thus possible that—e.g., as [Akresh et al. \(2013\)](#) find—a CCT would have a large impact on girls than on boys. For example,

parents may have already been planning to enroll boys and encourage their attendance at school, but only enrolled girls and encouraged their attendance in order to meet the conditions of the CCT. Third, for two central outcomes we hypothesized would vary according to the baseline quality of health facilities—health clinic visits and health during the last month—we examine heterogeneous impacts of the CCT program by baseline (2009) health clinic staff per capita. Specifically, we divide villages into two types: those with above-median and below-median healthcare staff per capita at baseline, where staff included doctors, nurses, and all other medical assistants and staff. This provides insight into how a village’s initial human resource capacity to address healthcare needs might mediate the impacts of a CCT program conditioned on health clinic visits. Of course, such results come with the caveat that healthcare personnel per capita may serve as a proxy for other important correlates of health-related outcomes—like overall poverty, or remoteness of the village—which could influence the interpretation of the results. In general, however, one would expect that villages with greater staffing levels would be better prepared to absorb potentially higher demand for healthcare, and therefore may see greater improvements in clinic visits and health due to the CCT program.

Fourth, for outcomes heavily related to demand for health and healthcare—whether one owns shoes or slippers, expenditure on insurance, participation in the CHF, whether one treats illness when it strikes and where (public or private facilities), and how one finances treatment received for illness—we examine heterogeneous impacts of the CCT program by baseline (2009) household poverty, as captured by asset wealth.⁹ While all beneficiary households are poor—on average, only 23 percent of village households are beneficiaries—this allows us to observe whether the moderately poor (top half of beneficiaries in terms of asset wealth) or the extremely poor (bottom half) benefit most from the program, and on which dimensions. One might expect the extremely poor—being the most vulnerable—to see the most health benefits from a CCT program (Akresh et al., 2013). To capture asset wealth and divide households into two groups, we carried out a principal components analysis (PCA) using dummy variables for ownership of 13 assets.¹⁰

Finally, hypothesizing that exposure to a negative shock would have a psychological impact on individuals and wipe out savings, we examine heterogeneous impacts of the CCT program according to exposure to a drought or flood shock within the five years preceding the baseline survey. The outcomes for which we examined impacts by shock exposure include livestock ownership, savings, and credit.

7.1 Health

In each of the three survey rounds, we collected individual-level data on health clinic visits, ownership of protective footwear (shoes and slippers) by children, 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 up hill, bending over or stooping, walking more than 1 km, walking more than 100 meters,

⁹While poverty is a multi-faceted, complex phenomenon, asset wealth provides a quick and objective insight into a household’s poverty status. Further, our survey did not capture detailed income data.

¹⁰These 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.

or using a bath or toilet), and 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). Data 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. In each round we also collected household-level data on expenditure on formal insurance,¹¹ and on whether the household participates in a government-run health insurance program known as the Community Health Fund (CHF). We examine the program impacts on all of these outcomes.

A few words about the CHF are warranted. In the early 1990s, the Tanzanian government introduced a health insurance program called the CHF. This program is a voluntary, district-level prepayment scheme. Members pay a fixed annual fee of between 5,000 and 10,000 Tanzanian shillings (between \$3 and \$6 US¹²), depending on the region, but then their entire family is exempt from any co-payments for visits to primary healthcare facilities (Marriott, 2011).¹³ Upon introduction of the CHF, children and maternal health services were already exempt from co-payments according to official government policy (Babbel, 2012). Ten years after the introduction of the CHF, the program had an average enrollment rate of only ten percent. At least two of the reasons cited for lack of participation were inability to pay or to see the rationale to insure (Kamuzora and Gilson, 2007). Insofar as liquidity has been a binding constraint, a cash transfer program might be expected to significantly impact participation.

7.1.1 Health clinic visits

In Table 4, we document the impact of the CCT program on the frequency of health clinic visits. We focus on overall impacts, impacts on children aged 0–5, and impacts on those age 60 and over, as the latter two groups are those for whom health conditions applied. (Furthermore, at midline, data on clinic visits were only gathered for those two age groups.) At midline (1.5 years after treatment began), treatment was associated with 2.3 more visits per year for children aged 0–5 (column 2) and 1.1 more visits per year for those aged 60 and over (column 3). These statistically significant effects, however, disappear at endline (2.5 years after treatment began) for both age groups. Strikingly, the baseline mean for both age groups exceeded the number of clinic visits required by the program conditions. Children aged 0–5 already visited health clinics 8.3 times per year on average at baseline, compared to the condition of 6 visits (59 percent of children already met this condition). Elderly aged 60 and older already visited health clinics 2.8 times per year on average at baseline, compared to the condition of 1 visit (65 percent of elderly already met this condition). Thus, the program’s emphasis on clinic visits may have led households to initially increase visits even though the average household was already satisfying the condition. Subsequently—by the endline—households’ 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. Another possibility is that health improvements due to the program that were realized by endline but not at midline reduced demand for clinic visits by endline.

¹¹Data on expenditure on insurance is unfortunately not further disaggregated by type of insurance.

¹²In 2009 the exchange rate ranged from 1,280 to 1,467 per U.S. dollar (Bank of Tanzania, 2015).

¹³Up to 7 family members are exempt from co-payments—though receipt of medications/ tests incurs fees.

Table 4: Effects of treatment on clinic visits

	Clinic visits		
	Full sample (1)	0-5 years old (2)	60 and over (3)
<i>Panel A: Effect of assignment to treatment</i>			
Treatment × 2011		2.296** (0.872)	1.083*** (0.349)
Treatment × 2012	-0.067 (0.253)	-1.042 (0.875)	0.161 (0.344)
2011		-3.817*** (0.584)	-1.214*** (0.216)
2012	-1.436*** (0.182)	-5.762*** (0.635)	-0.670*** (0.237)
Observations	13,713	1,243	5,692
R-squared	0.061	0.375	0.018
Baseline mean	2.802	8.272	2.783
<i>Panel B: Heterogeneous treatment effects by staff per capita</i>			
Treatment × 2011 × fewer		2.975*** (1.036)	0.875* (0.477)
Treatment × 2011 × more		1.426 (1.509)	1.279** (0.511)
Treatment × 2012 × fewer	-0.001 (0.373)	-1.483 (1.114)	0.199 (0.511)
Treatment × 2012 × more	-0.144 (0.331)	-0.234 (1.223)	0.126 (0.466)
Observations	13,713	1,243	5,692
R-squared	0.061	0.384	0.019
p-value of difference (midline)		0.400	0.565
p-value of difference (endline)	0.775	0.453	0.916

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. 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 $p < 0.10$.

Despite differences in the point estimates, we cannot reject the null hypothesis that the effects of the CCT program on health clinic visits in villages with few baseline health staff per capita (the bottom half of the distribution) are the same as those in villages with many health staff per capita (the top half). This is true overall and for both age groups. Given that staff per capita may reflect clinics' underlying ability to absorb increased demand, this is important; it suggests that the impacts of treatment on clinic visits would not be enhanced by increasing clinic staff per capita.

7.1.2 Health investments

While health clinic visits are an important aspect of individual investment in health, also important are purchases individuals make to prevent health problems from occurring or to cope with the risks posed by health shocks. Accordingly, we also examined whether the CCT program impacted take-up of three health-related products: shoes, slippers (i.e., open-toed footwear), insurance of any kind (e.g., medical or life), and government health insurance in particular (Table 5). Already by midline, the CCT program led to a significant, 18 percentage point increase in shoe ownership among 0–18 year old children that persisted until endline (column 1). A null impact on slipper ownership at midline changed to a significant, 8 percentage point increase by endline (column 2). This suggests that the CCT 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. This is important in light of evidence in the public health literature showing that use of shoes is associated with lower exposure to helminths ([Mascarini-Serra et al., 2011](#); [Birn and Solórzano, 1999](#)), and may help explain health improvements despite no sustained increases in the number of health clinic visits.

We also find increased household expenditures on insurance; the program led to a 6 fold increase in expenditure on insurance by midline, and a 8 fold increase by endline (column 3). One of the most striking and consistent impacts of the program, found for all sub-groups, is an increase in participation in Tanzania’s government-run health insurance program, the CHF. While we lack baseline data on participation rates in the CHF, at baseline only 3 percent of individuals who were sick during the last month and sought treatment reported using health insurance to fund that treatment. Our empirical analysis shows that by endline, the program made households in treatment villages 36 percentage points more likely to participate in the CHF (column 4). Given that access to health insurance allows individuals to more readily go to a health clinic when they are sick—rather than just when they have the money to do so—may be associated with better-timed clinic visits, and may therefore itself help explain health improvements.

The impacts of the CCT program on these four health investments are larger for the extremely poor (those among program beneficiaries who were in the bottom half of asset wealth at baseline) than for the moderately poor in each of the two survey rounds (Table 5, Panel B). 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 is slightly larger (but not statistically significantly different) for extremely poor households. Note that this does not merely reflect an income effect that boosted expenditures on all goods; as shown in another study, this CCT program did not increase food consumption (15). However, households did use the transfers to insure against health shocks. 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.

The qualitative analysis also found clear evidence of a significant impact of the CCT program on children's ownership of shoes. Every head teacher in the focus groups and in-depth interviews across districts stated that more of the beneficiary children were able to own school materials as a result of the program, including notebooks, uniforms, and shoes. Many teachers also emphasized that students need shoes to go to school: One teacher explained that while teachers may be lenient in the first few schools day of the year, students will very soon be turned away from school if they do not wear shoes.

The teachers' observation that more children owned shoes corresponds with the quantitative finding that the poorest children are relatively more likely to increase their ownership of shoes as a result of the program. These would have been the children that were most likely to have been unable to purchase shoes prior to the start of the CCT. The qualitative exercise did not find any evidence of significant changes in student behavior and activities outside of school.

A major finding of the qualitative fieldwork in almost every village visited was that many of the beneficiaries and community leaders involved stressed the importance of the community health fund (CHF) for beneficiary households. The heads of the health facilities across the villages involved in the qualitative work said that most or nearly all beneficiary households were enrolled in the program. In a village in Kibaha, the head of the dispensary said that when TASAF transfers are distributed, she sends the dispensary staff to the distribution point to sign up any household that is not yet participating or needs to renew their membership while they feel relatively "rich." The community management committee (CMC) in a village in Bagamoyo explained that it is important to "sensitize" the beneficiaries on how they should spend the money, and that they stress the importance of contributing to the CHF. This highlights community leader and service provider perceptions that the CHF is complementary to the CCT program, and that greater use of health insurance should be encouraged.

Table 5: Effects of treatment on 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 assignment to treatment</i>				
Treatment × 2011	0.180*** (0.043)	0.054 (0.035)	1.176*** (0.252)	
Treatment × 2012	0.179*** (0.047)	0.084** (0.038)	1.516*** (0.284)	0.357*** (0.039)
2011	0.129*** (0.028)	0.188*** (0.023)	0.177*** (0.051)	
2012	0.126*** (0.031)	0.196*** (0.028)	0.438*** (0.099)	
Observations	6,847	6,847	5,036	1,555
R-squared	0.105	0.107	0.118	0.317
Baseline mean	0.423	0.632	0.181	
<i>Panel B: Heterogeneous treatment effects by degree of poverty</i>				
Treatment × 2011 × extremely poor	0.300*** (0.065)	0.095* (0.051)	1.398*** (0.291)	
Treatment × 2011 × moderately poor	0.079* (0.045)	0.015 (0.035)	0.871*** (0.254)	
Treatment × 2012 × extremely poor	0.246*** (0.067)	0.159*** (0.049)	1.753*** (0.308)	0.380*** (0.050)
Treatment × 2012 × moderately poor	0.120** (0.055)	0.009 (0.046)	1.219*** (0.378)	0.328*** (0.040)
Observations	6,847	6,847	5,035	1,555
R-squared	0.110	0.112	0.122	0.318
p-value of difference (midline)	0.002	0.138	0.059	
p-value of difference (endline)	0.106	0.011	0.174	0.295

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Shoe and slipper ownership are individual-level outcomes for those 0-18 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.

7.1.3 Health and activities

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 (result available upon request), 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 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, but insignificant after using the [Benjamini et al. \(2006\)](#) method of correcting for multiple hypothesis testing) 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. 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.

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)
<i>Panel A: Effect of assignment to treatment</i>						
Treatment × 2011	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	-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	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	0.078*** (0.016)	0.031 (0.047)	0.147*** (0.023)	1.076*** (0.147)	0.298 (0.297)	2.389*** (0.269)
Observations	20,741	1,537	5,694	20,740	1,537	5,693
R-squared	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
<i>Panel B: Heterogeneous treatment effects by staff/capita</i>						
Treatment × 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)
Treatment × 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)
Treatment × 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)
Treatment × 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	20,741	1,537	5,694	20,740	1,537	5,693
R-squared	0.007	0.017	0.025	0.013	0.017	0.036
p-value of difference (midline)	0.668	0.926	0.474	0.197	0.901	0.414
p-value of difference (endline)	0.581	0.990	0.687	0.016	0.397	0.004

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. 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 $p < 0.10$.

We further explored the endline health impacts attributable to the CCT program by examining heterogeneous impacts by village health clinic staff per capita at baseline (Table 6, Panel B). Here, we find that reductions in sick days in the sample overall are concentrated in villages with more health staff per capita, with no significant impacts on number of sick days in villages in the bottom half of health staff per capita. For individuals in villages that were initially highly-staffed, the average reduction in sick days in the last month is 0.96 (compared to an insignificant 0.07 days for those in less highly-staffed villages). The difference between the effect of the program on sick days in highly-staffed vs. less highly-staffed villages is statistically significant at the 5 percent level. This suggests that reductions in the intensive margin of illness may in fact be conditional on

Table 7: Effects of treatment on activities of daily living

	Dummy - can ...						Ordinary activities index (7)
	Do vigorous activity (1)	Walk uphill (2)	Bend over or stoop (3)	Walk more than 1km (4)	Walk more than 100m (5)	Use bath or toilet (6)	
Treatment × 2011	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	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	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	-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)
Observations	5,685	5,685	5,685	5,685	5,403	5,685	5,403
R-squared	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

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Activity index is the sum of the six activity dummies; its range is 0 to 6. 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. 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 $p < 0.10$.

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, with fewer average annual visits at baseline than for young children. A further caveat is that there are no differential impacts of treatment on the incidence (extensive margin) of illness by clinic staffing levels. Clinic staffing appears to matter more for the extensive than the intensive margin of illness.

Despite overall health improvements, the CCT program did not change the ordinary activities that elderly individuals were able to perform, as shown in Table 7. 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 and becomes insignificant after adjusting for multiple hypothesis testing); there, we find a very small negative impact of the program that is statistically significant at endline. Overall, however, it is clear that 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.

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, but rather only after 2.5 years of transfers have been received.

7.1.4 Anthropometrics

Table 8 reports the effects of being in the treatment group 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 do not see statistically significant impacts of treatment on any of these outcomes—in either period. The lack of anthropometric effects is striking given the large number of metrics we consider; it contributes to a mixed literature on the impacts of CCTs on child anthropometrics ([Fiszbein and Schady, 2009](#)). This result is less surprising when considering other analysis from this program, showing that households did not use their transfers to increase food consumption (Table 15). Further, while expenditures on children’s protective footwear and insurance are likely to improve child health, it is possible that health benefits stemming from such investments take more time to materialize.

Health providers, in interviews and focus groups, did not report any measurable differences between beneficiary and nonbeneficiary children in terms of health or growth. However, educators in focus groups in Bagamoyo did state that children are now getting more to eat and so are better able to focus in school. Also, there could be considerable variation in anthropometric impacts, which is difficult to discern at the community level, or may not have been present in the specific communities visited as part of the endline focus groups and in-depth interviews.

Table 8: Effects of treatment on anthropometrics for children aged 0-5

	Z - scores						
	Height-for-age	Weight-for-age	Weight-for-height	BMI-for-age	Height (cm)	Weight (kg)	MUAC (mm)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment × 2011	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	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	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	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)
Observations	1,184	1,204	1,079	1,073	1,240	1,403	1,234
R-squared	0.081	0.087	0.075	0.073	0.065	0.074	0.037
Baseline mean	-1.354	-0.812	0.0519	0.197	87.24	12.15	155.8

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. 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. 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 $p < 0.10$.

7.2 Child Education and Activities

7.2.1 Education

In each of the three survey rounds, we collected individual-level data on literacy, whether children had ever attended school, their current school enrollment status, whether they passed the last national exam they took, whether they had missed school in the last week, and their activities in the last week (specifically, whether they had fetched water, cut wood, cleaned the toilet, cooked, cared for children, cared for the elderly, or received tutoring).

Panel A of table 9 shows the impacts of treatment on a variety of education-related outcomes for children 5–18 years old. We find that treatment was associated with a 4 percentage points (significant at the 10 percent level) increase in literacy (column 1) at midline — a 7 percent increase relative to the mean literacy rate. This effect, however, became insignificant by endline. Children in treatment households were also more likely to have attended school at some point. Treatment was associated with a 7 percentage point increase in having ever attended school (column 2) at midline and a 6 percentage point increase at endline. In the baseline survey, 76 percent of children aged 0–18 had ever attended school; thus, treatment was associated a 7–9 percent increase over the mean rate of children having ever attended school. Among children who had ever attended school, treatment did not significantly impact current school enrollment in either survey round (column 3).

Treatment was also not significantly associated with a child's parents reporting that they passed the last national exam for which they sat (here, we consider only children who completed Standard IV or higher). And finally, treatment was not associated with a lower likelihood of having missing school sometime in the last week.

Table 9: Effects of treatment on education

	Dummy - ...				
	Literate (1)	Ever attended school (2)	Currently in school (3)	Passed last national exam taken (4)	Missed school in last week (5)
<i>Panel A: Effects of assignment to treatment</i>					
Treatment × 2011	0.041* (0.024)	0.069*** (0.024)	-0.024 (0.027)	-0.038 (0.044)	0.022 (0.021)
Treatment × 2012	0.034 (0.027)	0.060** (0.026)	-0.025 (0.030)	-0.071 (0.043)	0.038 (0.025)
2011	0.198*** (0.017)	0.096*** (0.014)	-0.145*** (0.021)	-0.141*** (0.023)	-0.034** (0.015)
2012	0.247*** (0.019)	0.122*** (0.016)	-0.215*** (0.022)	-0.105*** (0.025)	-0.070*** (0.014)
Observations	5,460	5,460	4,625	2,241	5,448
R-squared	0.182	0.112	0.146	0.073	0.014
Baseline mean	0.568	0.763	0.883	0.821	0.0808
<i>Panel B: Heterogeneous treatment effects by gender</i>					
Treatment × 2011 × girl	0.056* (0.031)	0.078*** (0.029)	-0.022 (0.035)	-0.047 (0.065)	0.030 (0.027)
Treatment × 2011 × boy	0.026 (0.034)	0.060* (0.030)	-0.026 (0.035)	-0.031 (0.063)	0.015 (0.026)
Treatment × 2012 × girl	0.063* (0.035)	0.072** (0.032)	-0.039 (0.038)	-0.111 (0.073)	0.048 (0.029)
Treatment × 2012 × boy	0.010 (0.036)	0.049 (0.031)	-0.013 (0.040)	-0.035 (0.058)	0.029 (0.029)
Observations	5,460	5,460	4,625	2,241	5,448
R-squared	0.183	0.112	0.146	0.073	0.014
p-value of difference (midline)	0.506	0.613	0.925	0.863	0.645
p-value of difference (endline)	0.269	0.497	0.591	0.440	0.548

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. All outcomes are restricted to those 5 - 18 years old in the baseline survey. The universe for the *in school* outcome (column 3) is youth who have ever attended school. The universe for the *national exam* outcome (column 4) is youth who have at least Standard 4 education and who have taken a national exam. The youngest youth in our data who had taken a national exam was 8 years old. Missed school (column 5) excludes missing school for reasons such as school being closed or teacher being absent. 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 $p < 0.10$.

Panel B of table 9 explores heterogeneous effects of treatment by gender. Treatment's effects on both literacy and having attended school are generally larger in magnitude for girls than for boys.

For example, whereas in the pooled estimates of Panel A the endline treatment effect on literacy is insignificant, the endline treatment effect for girls is statistically significant. Treatment is associated with 5.6 and 6.3 percentage points increase in literacy among girls at midline and at endline, respectively (equivalent to a 10 percent and an 11 percent increase over mean literacy rates of girls, respectively). However the differences between the boys and girls, for this outcome and all others in this table, are statistically insignificant. Like literacy, treatment effects on having attended school are also stronger for girls. Like the pooled estimates across boys and girls, treatment is not significantly associated with current school enrollment, passing last national exam, or missing school.

One village leader conceded that the program may not have increased enrollment, but that it did increase their attentiveness and confidence when they did attend. The qualitative work further clarified that previously, some children had been unable to pay attention due to lack of food, and that they were ashamed of not having the appropriate uniform, but that morale improved after the program began.

The qualitative exercise also found a range of responses from communities on how they encouraged children to attend school. In one village in Kibaha, the village executive officer (VEO) told us that he would receive information from the CMC and school, and if some children were not complying with the conditions of the CCT program, he would contact the family directly. This VEO told us that while the CMC lacked power to induce households to change their behavior, his involvement was enough to ensure that children attended school. However, in other communities village leaders expressed less engagement with monitoring the conditions of the program.

7.2.2 Child Activities

Table 10 shows that treatment had few statistically significant impacts on the likelihood that a child would take part in a variety of children's activities. The only activity included in the survey that was significantly affected by treatment was toilet cleaning; it was significantly reduced at midline, but this effect disappeared by endline. The other six activities were unaffected. Overall, we conclude that the CCT had little impact on what children did with their days.

Table 10: Effects of treatment on childrens' activities

	Dummy - child has ... in last week						
	Fetches water (1)	Cut wood (2)	Cleaned toilet (3)	Cooked (4)	Cared for children (5)	Cared for elderly (6)	Received tutoring (7)
Treatment × 2011	-0.004 (0.046)	-0.042 (0.049)	-0.056* (0.029)	-0.019 (0.037)	-0.003 (0.037)	0.013 (0.049)	0.007 (0.013)
Treatment × 2012	0.005 (0.048)	0.033 (0.036)	-0.021 (0.035)	-0.002 (0.031)	-0.012 (0.034)	-0.003 (0.047)	-0.007 (0.013)
2011	0.236*** (0.033)	0.333*** (0.034)	0.106*** (0.015)	0.261*** (0.026)	-0.022 (0.028)	-0.105*** (0.038)	-0.004 (0.009)
2012	0.127*** (0.030)	0.173*** (0.023)	0.101*** (0.023)	0.174*** (0.022)	-0.057** (0.022)	-0.130*** (0.037)	0.023** (0.010)
Observations	5,477	5,477	5,477	5,477	5,476	5,477	5,476
R-squared	0.078	0.126	0.023	0.119	0.010	0.034	0.007
Baseline mean	0.571	0.280	0.131	0.259	0.163	0.282	0.021

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. All outcomes are restricted to those 4 - 18 years old in the baseline survey. In Tanzania, *tuition* refers to receiving tutoring (often paid for) outside of school. All specifications include individual 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 $p < 0.10$.

7.3 Household Assets

In each of the three survey rounds, we collected household-level data on the number of acres of land owned, 15 durable household assets (iron, refrigerator, television, mattress or bed, radio, watch or clock, sewing machine, stove (excluding traditional stove), bicycle, motorcycle, car or truck, wheelbarrow, mobile phone, sponged sofa, non-sponged sofa) and ten livestock assets (dairy cows (including calves), indigenous cows (including calves), dairy goats (including baby goats), indigenous goats (including baby goats), chickens (local variety, excluding chicks), chickens (foreign variety, excluding chicks), sheep, pigs, ducks/turkeys, and rabbits). We report treatment effects on ownership of nine of the 15 durable household assets (the six assets that we omit were present in less than one percent of the households at baseline) and all livestock except for dairy cows and rabbits (no household owned either animal at baseline).

In Table 11, we document the impact of the CCT program on acres of land owned and on indicators for ownership of nine durable household assets. Treatment is associated with almost a 6 percentage point increase in bike ownership at midline (significant at the five percent level). Since 19 percent of households owned a bike at baseline, this represents a 30 percent increase relative to the mean. However, this effect dissipates by endline. There were no other significant impacts of treatment on durable household asset ownership.

One possible explanation for the lack of significant treatment effects on durable household asset ownership is that the types of assets purchased were too diverse to register in the analysis of individual items. For example, one man in a focus group in Bagamoyo described how he used the transfer income to purchase a hammer, which he uses to make gravel that he sells for additional income. While this reflects a productive investment in a durable asset, the categories in the quantitative data may not capture every diverse investment.

One focus group of beneficiaries in Bagamoyo explained that some of the elderly are now able to pay people to cultivate their land, which allows them to generate additional income. Several elderly beneficiaries in different villages in Kibaha also said that they used the money to clear land or assist with farming. Therefore, while the program may not increase land ownership, it may serve to increase the returns beneficiaries are able to get from land they already own.

Table 11: Effects of treatment on household assets

	Dummy - household owns ...									
	Number of acres of land (1)	Mattress or bed (2)	Radio (3)	Bike (4)	Mobile phone (5)	Watch or clock (6)	Stove (7)	Iron (8)	Padded sofa (9)	Unpadded sofa (10)
Treatment × 2011	0.266 (0.325)	-0.036 (0.044)	0.032 (0.029)	0.057** (0.026)	0.020 (0.028)	-0.008 (0.017)	0.030 (0.020)	0.002 (0.013)	0.005 (0.010)	0.019 (0.020)
Treatment × 2012	-0.192 (0.330)	0.017 (0.036)	0.033 (0.033)	0.015 (0.027)	-0.021 (0.028)	-0.004 (0.016)	0.022 (0.023)	0.001 (0.013)	-0.006 (0.011)	0.008 (0.011)
2011	-0.075 (0.190)	0.100*** (0.026)	0.016 (0.018)	-0.004 (0.020)	0.123*** (0.020)	-0.004 (0.011)	0.020 (0.016)	0.009 (0.010)	0.009 (0.008)	0.043*** (0.014)
2012	-0.343** (0.169)	-0.001 (0.024)	-0.050** (0.022)	0.009 (0.019)	0.191*** (0.018)	-0.037*** (0.011)	0.047*** (0.016)	0.010 (0.011)	0.020** (0.009)	0.015 (0.010)
Observations	5,023	5,036	5,036	5,036	5,036	5,036	5,036	5,036	5,036	5,036
R-squared	0.006	0.018	0.008	0.004	0.074	0.006	0.011	0.001	0.004	0.024
Baseline mean	4.039	0.713	0.327	0.190	0.105	0.103	0.097	0.051	0.023	0.012

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. In Tanzania, sometimes stones are arranged to hold a pot above a fire — but that is not considered a stove here. 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 $p < 0.10$.

Panel A of table 12 reports the effects of the CCT program on household livestock ownership. Treatment households significantly increased the number of goats and chickens (local variety) they owned. Treatment was significantly associated with households owning 0.4 more indigenous goats in both the midline and endline surveys. At midline, treatment resulted in owning about 0.9 more local chickens; this effect grew to 1.0 by endline (both significant at the 1 percent level). Ownership of other livestock—indigenous cows (including calves), dairy goats (including baby goats), indigenous goats (including baby goats), chickens (foreign variety, excluding chicks), ducks/turkeys, sheep, and pigs—was largely unaffected by the CCT.

In panel B of table 12 we explore the heterogeneous impacts of the CCT program by whether the household had reported experiencing a drought or flood shock during the five years prior to the baseline survey. Treatment had a remarkably consistent effect on goat ownership by this shock. At both midline and endline, treatment was associated with 0.3 more goats for households that had experienced this shock but 0.5 goats for those that had not (column 3). This difference, however, is not statistically significant. Local chickens also followed this pattern (column 4), with chicken ownership increasing more at midline in households without this weather shock than in those with it. The difference here is statistically significant at the 10 percent level. At endline, interestingly, the pattern reverses and the CCT's effects on chicken ownership are stronger for the households that had experienced a drought or flood. At midline, treatment is also associated with more sheep ownership in households that have experienced a drought/flood. It is difficult to understand why different types of livestock were affected differently by the CCT based on whether the household had experienced this weather shock. Perhaps certain livestock have real or perceived advantages to dealing with weather shocks.

Consistent with the quantitative findings of significant, large increases in goat and chicken ownership, almost every beneficiary participant in the focus group and in-depth interview discussions mentioned that they had bought chickens, goats, or even ducks with the transfer money. This use of the transfer money seemed to be widespread knowledge in the villages: even one focus group of nonbeneficiaries in Bagamoyo explained that the beneficiaries used the transfer money to invest in chickens. This seemed to be the most prevalent store of value in these villages, and functions as a type of savings for these vulnerable households. For example, one old man in a focus group in Bagamoyo said that he had used the money from the transfers to purchase a chicken, which he then sold to have money to pay for someone to cultivate his land. Another woman in Bagamoyo mentioned that she had created a business of cooking and selling the meat from chickens she was able to buy with the CCT transfers. A man in Kibaha described how the transfer money had allowed him to buy chickens and ducks, which had then reproduced so that he could sell the chicks and ducklings for income.

Table 12: Effects of treatment on livestock ownership

	Number of ... household owns today							
	Indigenous cows	Dairy goats	Indigenous goats	Local variety chickens	Foreign variety chickens	Sheep	Pigs	Turkeys and ducks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Effect of assignment to treatment</i>								
Treatment × 2011	-0.023 (0.077)	0.026 (0.027)	0.401*** (0.130)	0.918*** (0.311)	0.211 (0.219)	0.022 (0.046)	0.009 (0.024)	-0.016 (0.051)
Treatment × 2012	0.002 (0.135)	0.020 (0.025)	0.404** (0.161)	0.993*** (0.349)	-0.055 (0.040)	-0.022 (0.033)	-0.016 (0.019)	-0.060 (0.069)
2011	0.050** (0.023)	0.015 (0.014)	0.027 (0.045)	-0.366* (0.196)	0.029* (0.016)	0.032 (0.036)	-0.002 (0.019)	0.029 (0.027)
2012	0.144* (0.081)	0.014 (0.015)	0.106* (0.061)	-0.109 (0.192)	0.067** (0.031)	0.065*** (0.021)	0.006 (0.017)	0.086* (0.047)
Observations	5,035	5,035	5,035	5,034	5,035	5,035	5,035	5,035
R-squared	0.002	0.002	0.011	0.008	0.001	0.004	0.001	0.001
Baseline mean	0.0790	0.001	0.252	2.367	0.102	0.022	0.019	0.136
<i>Panel B: Heterogeneous treatment effects by drought or flood shock</i>								
Treatment × 2011 × no drought/flood	0.013 (0.111)	0.003 (0.003)	0.539* (0.286)	1.529*** (0.482)	0.016 (0.059)	-0.058 (0.093)	0.044 (0.053)	0.036 (0.081)
Treatment × 2011 × drought/flood	-0.045 (0.104)	0.040 (0.043)	0.311** (0.126)	0.522 (0.351)	0.338 (0.358)	0.075* (0.045)	-0.014 (0.018)	-0.049 (0.062)
Treatment × 2012 × no drought/flood	0.058 (0.275)	0.026 (0.018)	0.510* (0.289)	0.626 (0.461)	-0.104 (0.071)	-0.045 (0.057)	-0.003 (0.037)	-0.099 (0.119)
Treatment × 2012 × drought/flood	-0.032 (0.119)	0.016 (0.034)	0.334* (0.189)	1.223*** (0.412)	-0.020 (0.053)	-0.007 (0.036)	-0.025 (0.018)	-0.037 (0.081)
Observations	5,034	5,034	5,034	5,033	5,034	5,034	5,034	5,034
R-squared	0.003	0.004	0.011	0.010	0.002	0.005	0.002	0.002
p-value of difference (midline)	0.701	0.401	0.483	0.066	0.378	0.201	0.283	0.389
p-value of difference (endline)	0.758	0.728	0.614	0.257	0.372	0.547	0.575	0.664

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Young animals included also, except for chicks. Drought or flood refers to households reporting that they experienced a drought or flood in the 5 years prior to the baseline survey. All specifications include household fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

7.4 Expenditures and Finance

In each of the three survey rounds, we collected household-level savings and credit outcomes (whether someone in the household has a bank account, has non-bank savings, has taken out a loan in the last year) non-food expenditures (tobacco products, children’s clothing, adult clothing, personal effects, weddings/funerals/dowries, medical services, medication, and boarding school costs), and food consumption (maize, rice, beans, and sugar).

7.4.1 *Savings and credit*

Panel A of table 13 reports the effect of the CCT on three savings and credit outcomes. We see that treatment had no significant impacts on the likelihood that a household had a bank account, had other (non-bank) savings, or had borrowed from someone outside the household or from an institution within the past year. At baseline, less than two percent of households had a bank account, and about one percent had non-bank savings.

When we look at the heterogeneous treatment effects by drought or flood (13, Panel B), we still see no effects on bank account ownership, though treatment did increase non-bank savings. These increases in non-bank savings, however, were only observable in households that had not experienced a flood/drought in the five years preceding the baseline survey. At midline, treatment was associated with being 4.2 percentage points more likely to have non-bank savings. This effect increased to 6.4 percentage points at endline. These are large effects; at baseline only 1.7 percent of households that had not experienced a flood or shock had other savings. The difference between estimates for the two groups (exposed and unexposed to a flood/drought shock) is significant at endline, but not at midline. The CCT also increased borrowing for households that had not experienced a flood/drought at midline, but the effect disappeared by endline. A CCT increasing both borrowing and having savings is not necessarily a contradiction — households with savings are more likely to be given a loan.

The qualitative fieldwork supports the lack of significant findings of treatment on savings. Most beneficiaries in both focus groups and in-depth interviews said that money from CCT transfers was enough to pay their necessary expenses but not more, or even that they would often run out of money before the next transfer. Thus, while some households may have been able to increase their nonbank savings, this was not common across treatment households, and is consistent with the quantitative findings of a positive but insignificant effect on nonbank savings overall.

Various focus groups and interviewees reported a range of contrasting impacts of the program on their likelihood of borrowing money. One focus group of community leaders in Bagamoyo district described how the community as a whole had reduced borrowing, since there was now more money in circulation. One beneficiary in Bagamoyo district and another in Kibaha said that they still needed to borrow money to cover expenses before receiving a transfer, although they were able to pay off these debts as soon as they received the money.

Table 13: Effects of treatment on household savings and credit

	Dummy - someone in the household ...		
	Has a bank account (1)	Has other savings (2)	Borrowed in past year (3)
<i>Panel A: Effect of assignment to treatment</i>			
Treatment × 2011	-0.008 (0.008)	0.024 (0.015)	0.038 (0.030)
Treatment × 2012	-0.002 (0.009)	0.024 (0.019)	0.001 (0.036)
2011	0.009 (0.006)	0.025*** (0.009)	0.011 (0.022)
2012	0.004 (0.006)	0.056*** (0.013)	0.048* (0.025)
Observations	5,035	5,033	5,035
R-squared	0.001	0.028	0.005
Baseline mean	0.018	0.012	0.188
<i>Panel B: Heterogeneous treatment effects by drought or flood shock</i>			
Treatment × 2011 × no drought/flood	-0.002 (0.011)	0.042** (0.017)	0.068* (0.028)
Treatment × 2011 × drought/flood	-0.011 (0.012)	0.012 (0.020)	0.019 (0.039)
Treatment × 2012 × no drought/flood	0.004 (0.013)	0.064* (0.032)	0.012 (0.047)
Treatment × 2012 × drought/flood	-0.006 (0.013)	-0.002 (0.020)	-0.005 (0.047)
Observations	5,034	5,032	5,034
R-squared	0.001	0.029	0.006
p-value of difference (midline)	0.603	0.188	0.322
p-value of difference (endline)	0.549	0.059	0.775

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Borrow refers to borrowing from someone outside the household or from an institution receiving either cash, goods or services. Drought or flood refers to households reporting that they experienced a drought or flood in the 5 years prior to the baseline survey. All specifications include household fixed effects. Standard errors are in parentheses and clustered at the village level. *** indicates $p < 0.01$; ** indicates $p < 0.05$; and * indicates $p < 0.10$.

7.4.2 Non-food expenditures

We report the CCT's effects on annual non-food expenditures in table 14. Treatment was associated with spending an additional 3,228 Tsh on children's clothing (column 2) annually at midline (significant at the 5 percent level). This effect became insignificant by the time of the endline survey. Treatment also increased expenditures on other personal effects (431 Tsh) and medication (1,521 Tsh) — but only at midline. Treatment did not significantly affect expenditures on tobacco products, adult clothing, weddings/funerals/dowries, medical services, or boarding school.

This is consistent with the qualitative analysis. Almost every beneficiary focus group and in-

depth interview participant across districts stated that the money they received from the CCT program was spent on children's school supplies and on chickens and other livestock, with the remainder used to purchase food. However, except for indirect inclusion of school uniforms in the categories for men and boys clothing and women and girls clothing, these most common expenditures are largely absent from the standard nonfood expenditure categories considered in this section. Since there are few significant quantitative impacts of the program on nonfood expenditures at endline, the explanation may relate to the fact that since the transfers are relatively small, there was not enough money to use for additional purchases outside of the most common expenditures on livestock and school supplies.

The significant effect of treatment on medication expenditures is supported by the qualitative evidence. For example, one head of a health facility in Kibaha said that many more people are coming to the health facility that could not previously afford to visit. One man in Kibaha also said that the transfers had helped him to pay for medication when the dispensary was out of drugs.

Table 14: Effects of treatment on non-food expenditures

	Expenditures (TSH) on ... over last 12 months								
	Tobacco products (1)	Children's clothing (2)	Clothing/ footwear for men (3)	Clothing/ footwear for women (4)	Other personal effects (5)	Weddings/ funerals/ dowries (6)	Medical services (7)	Medication (8)	Boarding school costs (9)
Treatment × 2011	-1,103 (1,311)	3,228** (1,331)	174 (1,142)	1,749 (1,665)	431* (242)	687 (638)	2,938 (1,871)	1,521* (868)	2,582 (2,270)
Treatment × 2012	-1,242 (1,326)	1,774 (1,483)	-767 (1,124)	930 (1,588)	416 (273)	899 (755)	-799 (2,620)	387 (927)	2,237 (3,014)
2011	623 (923)	116 (914)	917 (792)	402 (1,164)	-501** (194)	-703 (493)	-4,397*** (1,346)	-2,593*** (716)	-3,340 (2,043)
2012	772 (768)	3,309*** (1,157)	2,437*** (804)	2,375* (1,229)	-443* (235)	721 (593)	3,371* (1,897)	451 (652)	-898 (2,610)
Observations	5,030	5,029	5,026	5,030	5,030	5,029	5,025	5,028	5,029
R-squared	0.001	0.020	0.006	0.008	0.004	0.007	0.011	0.018	0.002
Baseline mean	5265	6389	5971	8772	814	3934	9412	4984	4552

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Children refers to household members under 15 years old. Men/women refer to those at least 15 years old. Medical services and medication excludes traditional medicine. 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 15: Effects of treatment on food consumption

	Value of food consumption (TSH) on ... over last 7 days)									
	Maize (flour/super/sembe)		Maize (dona)		Other flour (millet/cassava/ sorghum/barley)		Husked rice		Dried beans	Sugar
	Purchased (1)	Produced (2)	Purchased (3)	Produced (4)	Purchased (5)	Produced (6)	Purchased (7)	Produced (8)	Purchased (9)	Purchased (10)
Treatment × 2011	62 (586)	714 (657)	-41 (213)	-102 (194)	-21 (65)	-397* (207)	-18 (280)	-359 (379)	24 (122)	-109 (140)
Treatment × 2012	-354 (413)	662 (840)	269 (197)	274 (255)	-17 (67)	45 (125)	-441 (335)	-152 (318)	19 (137)	133 (161)
2011	-2,110*** (396)	1,872*** (483)	-404*** (136)	350*** (108)	-30 (39)	357* (198)	1,016*** (214)	1,260*** (293)	857*** (98)	554*** (107)
2012	742*** (268)	2,885*** (583)	-358*** (128)	401** (156)	29 (42)	62 (89)	1,597*** (295)	793*** (263)	1,119*** (105)	992*** (118)
Observations	4,568	4,267	4,764	4,879	4,716	4,925	4,834	4,535	4,982	4,962
R-squared	0.113	0.151	0.020	0.020	0.002	0.011	0.065	0.063	0.195	0.104
Baseline mean	3600	216	670	371	145	424	263	61	551	580

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Production refers to made in the home (not purchased). The values of produced sugar and dried beans is not included as home production of these two goods is negligible. 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 $p < 0.10$.

7.4.3 Food consumption

Table 15 shows that the CCT had little effect on food consumption over the last week — whether purchased or produced at home (estimated value). The only significant (at the 10 percent level) effect was a reduction of 397 Tsh spent on produced “other” (millet/cassava/sorghum/barley) flour.¹⁴ Point estimates are large relative to baseline values, with large standard errors, so challenges in measurement may make it more difficult to estimate changes in food consumption precisely.

While there is little evidence of an increase in food consumption across diverse categories due to the CCT program, beneficiaries across focus groups and in-depth interviews in all three districts claimed that they had increased their food consumption as a part of the program. However, when asked for a breakdown of their expenditures with the last transfer, the food portion was generally small. As such, any increase in food consumption in beneficiary households relative to control households may be too small to outweigh any time, seasonality, and recall noise in the consumption data. Another possible explanation for the fact that so many people reported increases in food consumption is that consumption actually did increase. However, this occurred in both villages that received the program and those that did not, as shown by the fact that the “2012” coefficient is large and significant for most of the food items in table 15. While beneficiaries may attribute their increased food consumption to the program, increased food consumption occurred in control villages as well.

7.5 Community relations; data and outcomes

In each of the three survey rounds, we collected household-level information on whether the head of the household agrees with various trust statements and the value (in TSH) of transfers going in and out of the household.

7.6 Trust and Transfers

7.6.1 Trust

In Table 16, we document the impact of the CCT on self-reported trust in community members. Both at midline, treatment is associated with a 5.2 percentage point increase in the share of households reporting that leaders can generally be trusted (column 1). At endline, this effect has a similar magnitude but is more significant. These treatment effects represent a 6 to 7 percent increase over the baseline mean of 0.81. Unexpectedly, treatment is associated with a 6.2 percentage point decline in households’ reported trust in “most people” at midline and dissipates away at endline (column 2). Trust in community members was not significantly effected at either midline or endline.

The qualitative exercise found complex effects of the program on community dynamics and trust across villages. In some of the focus groups in Bagamoyo, beneficiaries stated that the program had made people feel that their leaders and community cared about them, and so improved

¹⁴The values of produced sugar and dried beans is not included since home production of these two goods is negligible.

community dynamics. This finding of increased trust in leaders corresponds to the quantitative findings of increased trust in treatment villages.

However, there were also reports from some leaders and nonbeneficiaries that the poorest households were not always selected. One man in Kibaha district told us that while some of the community leaders benefited from the program, he had not been chosen even though he did not have much money, and this may have been due to the fact that he was from a different tribe. We also heard from one CMC member in Kibaha and a focus group of nonbeneficiaries in Bagamoyo that the process had not been completely fair in those communities, and that some of the poorest households had been left out of the selection process. Even if based purely on perception rather than reality, such sentiments would tend to erode some of the trust within the community. This may account for the short-term reduction in trust of others in the community, captured in the midline survey. However, those effects disappeared on average by the time of the endline survey—possibly due to community members’ ability to observe the typical conditions of extreme poverty of those who were chosen to participate in the program.

Table 16: Effects of treatment on community trust

	dummy - ... can be trusted		
	Leaders (1)	Most people (2)	Community people (3)
Treatment × 2011	0.052* (0.031)	-0.062* (0.034)	-0.040 (0.030)
Treatment × 2012	0.054** (0.027)	0.016 (0.042)	0.028 (0.043)
2011	-0.027 (0.022)	0.301*** (0.026)	0.198*** (0.021)
2012	-0.052*** (0.016)	-0.106*** (0.028)	0.115*** (0.028)
Observations	5,007	4,996	4,993
R-squared	0.003	0.169	0.042
Baseline mean	0.805	0.242	0.556

Note: From authors’ calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. 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 $p < 0.10$.

7.6.2 Transfers

Table 17 documents the treatment’s effect on the value of transfers received by households, in panel A. Treatment households by definition receive a cash transfer from TASAF — part of the Tanzanian government— so it is unsurprising that treatment is associated with a large transfer from the government (column 3). Treatment was associated with transfers from the government that were 100,610 TSH at midline and 87,944 TSH at endline (significant at the 1 percent level). However, this large transfer seems to have a small and transitory “crowding-out” effect on community assistance. While at midline treatment is associated with receiving 11,841 TSH less from

individuals (an amount equivalent to about 12 percent of the value of the increase in transfers from government), this effect disappears by endline.

Table 17: Effects of treatment on transfers over last 12 months

Panel A: Effect of assignment to treatment on transfers into household

	Total transfers (TSH) from ...			
	Total	Individuals	Government or TASAF	NGOs
	(1)	(2)	(3)	(4)
Treatment × 2011	82,237*** (6,670)	-11,841** (5,936)	100,610*** (4,382)	-382 (480)
Treatment × 2012	76,925*** (7,559)	-4,535 (6,549)	87,944*** (3,578)	-413 (572)
2011	19,573*** (4,828)	24,322*** (5,356)	329 (588)	119 (375)
2012	19,416*** (4,945)	26,755*** (5,292)	-381*** (133)	253 (515)
Observations	5,370	5,032	5,032	5,025
R-squared	0.218	0.037	0.565	0.000
Baseline mean	21853	20229	648	942

Panel B: Effect of assignment to treatment on transfers out of household

	Value of ... gifts/assistance (TSH)			
	Total	Cash	Food	Other in-kind
	(1)	(2)	(3)	(4)
Treatment × 2011	74 (559)	195 (176)	-473 (380)	98 (134)
Treatment × 2012	-643 (607)	14 (234)	-868** (390)	199 (243)
2011	962** (469)	56 (112)	1,200*** (308)	-9 (101)
2012	1,948*** (479)	376** (182)	1,573*** (290)	339* (177)
Observations	5,028	5,029	5,026	5,028
R-squared	0.010	0.004	0.016	0.007
Baseline mean	1100	315	396	138

Note: From authors' calculations based on baseline (2009), midline (2011), and endline (2012) household survey data. Transfers into household include in-kind transfers. 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 $p < 0.10$.

In panel B of table 17 we document the value of gifts/assistance that households are giving. Households were selected for treatment partially based on need, so it is not surprising that the values in this table are relatively small. Total transfers out of the households were not significantly impacted, but treatment did reduce food assistance by 868 TSH at endline. This is a modest

reduction, equivalent to less than 1 percent of the overall value of government transfers received by the household.

The qualitative fieldwork supports the finding that in beneficiary households there may be occasional reductions in transfers from individuals. One nonbeneficiary in a focus group in Bagamoyo district said that she was very grateful for the transfers, since now she did not need to worry so much about supporting her parents.

In-depth interviews with village leaders in a treatment village in Kibaha revealed that a few of the elderly females had no other source of income aside from the program. Similarly, participants in a focus group for community leaders in Bagamoyo said that there are very few people in their village that receive money from outside. This reflects the quantitative findings of the small amount of all types of transfers before the program, relative to the much larger transfers received from the government as a result of the CCT program.

7.6.3 Heterogeneous effects of treatment by community scorecard

Since half of the treatment communities randomly were assigned to the CSC module, we also explored heterogeneous effects of treatment by participation in the scorecard activities (see 6.2: Monitoring System). Unfortunately, most of the scorecards' effects were insignificant. There are a total of 14 tables with treatment estimates (tables 4 - 17), and those tables contain 178 treatment estimates. By random chance alone one would expect about 18¹⁵ of the estimates to be significant. Just 16 of the estimates were significant ($p < 0.10$), which is completely consistent with what one would expect by spurious random chance, and accordingly we do not believe the CSC activities had any notable effects on the CCT's effects.

8 Conclusions and Policy Implications

This report describes the design and implementation of the impact evaluation of a community-managed conditional cash transfer program which began in January 2010. The evaluation instruments include household surveys carried out at baseline (late 2009), midline (mid 2011) and endline (late 2012); a community scorecards exercise carried out in 20 treatment communities in late 2010 — early 2011; two rounds of focus groups (in six villages following the midline survey and in nine villages following the endline); and a set of 39 in-depth interviews in six communities following the endline.

The baseline survey showed that the households in treatment and control communities are comparable across a broad range of characteristics: household size, access to financial services, household infrastructure, school enrollment, health seeking behavior, and involvement in community activities. The midline survey, carried out after 18-21 months of transfers had been realized, showed a range of significant impacts. Participating households were much more likely to visit health clinics in the previous year, especially for elderly members of those households. Illness rates for the previous month were not significantly lower, but children were more likely to be currently enrolled in school. The program also led households to enroll children in school at younger

¹⁵A p-value < 0.10 implies type 1 errors in in 10% of cases.

ages, and improved grade progression.

Significant impacts are observed across a broad array of areas, including health, education, and various risk-reducing behaviors: use of health insurance, insurance expenditures, non-bank savings (for the poorest households), and the purchase of livestock such as goats and chickens. In addition, the program has led to significant increases in spending on certain children's goods (especially children's shoes). We do not, however, find significant increases in food consumption. Improvements were not the same for all households and villages. In particular, health improvements (as measured by reductions in sick days) were largest in villages with more baseline health workers per capita, consistent with improvements being sensitive to capacity constraints.

Results on trust suggest that the program is associated with an increase in trust in community leaders as well as increased trust in some sub-groups of community members. Concerns about the potential adverse impact of the program on community trust thus seem unfounded. While some crowding out of individual transfers and a reduction in trust in the community were observed at midline, by endline there was no evidence of either. In fact, trust in community leaders and various communities members appears to have increased substantially. Increases in purchases of preventative health products—like insurance and shoes—were generally higher among the extremely poor (those in the bottom half of households on an asset index) than among the moderately poor. And increases in asset accumulation and savings due to treatment were somewhat greater in villages less exposed to droughts and floods as of baseline than to those more exposed.

Overall, the program has shown strong impacts in many areas: health outcomes, education outcomes, risk-reducing behaviors, and investments in livestock to improve livelihoods. This suggests that community-managed conditional cash transfers are a promising investment to reduce risk and improve human capital investments for the most vulnerable households in Tanzania.

The pilot CCT program underwent an initial scale-up extending the program to cover 275,000 households from all districts in the country, including Zanzibar. Now, a second scale-up has enrolled 1.1 million households with about 5 million beneficiaries across more than 9,900 villages in all 161 districts of Tanzania. The scale-up has come with a second round of impact evaluations, including an overall impact evaluation to examine whether the same effects observed at pilot are also observed at scale up, as well as narrower evaluations to examine the impact on women's empowerment and youth's well-being and transition to adulthood. This second-round scale up has been at least partially enabled through World Bank financing. In the project appraisal document for the US\$220 million project to support the scale-up, the pilot evaluation is referenced multiple times. For example, "The selected approach [for the project scale-up] also responds to recent analytical work on poverty and vulnerability in Tanzania, evaluations of the program and the pilot CCT." Targets for improvements in well-being in the project scale-up were drawn from the results of the pilot, and the results of the impact evaluation are featured, along with results of the process evaluation, a targeting assessment, and a community score card exercise ([WB2, 2012](#)).

The initial design of the pilot was determined over the course of a series of consultations between TASAF — the implementing agency — the World Bank, third-party groups with expertise in implementing cash transfer programs elsewhere in the world, and a range of stakeholders in Tanzania, including other government groups, communities, and non-government organizations.

The evaluation team worked closely with the government (TASAF) throughout the course of the evaluation but has not been directly involved in the scale-up. Rather, the evaluation team shared all impact evaluation results with TASAF, and TASAF worked with World Bank operational teams as well as other donors to design and secure resources for the scale-up. TASAF has reported that the results of the impact evaluation were key to obtaining buy-in, as well as more qualitative results. The expansion has been reported on in [Costa et al. \(2016\)](#):

"The expansion (both in the quantitative and spatial dimensions) was conceived as part of Tanzania's efforts to achieve Millennium Development Goal (MDG) 1: to halve the proportion of people living on less than USD1.25 per day. ... In this context, the government was successful in securing adequate funding for the scale-up, most notably through a World Bank credit of USD220 million (with the possibility of additional financing of USD200 million), the commitment of its own government resources, as well as the support of several development partners such as the United Kingdom Department for International Development (DFID), the Swedish International Development Cooperation (SIDA), the United States Agency for International Development (USAID) and the United Nations (UN). Finally, to achieve a successful expansion, the government has adopted operational adjustments, established a solid unified registry of beneficiaries, planned impact evaluations of the project and assembled teams to build capacity among local government PSSN staff on all aspects of the program, such as targeting, enrollment and compliance processes."

Program details during the scale-up differ from those during the pilot in six main ways. Three of those ways (items 3, 4, and 6 below) were directly informed by the findings of the pilot impact evaluation. The other three grew out of operational experience of the implementing government team.

1. While the targeting method is the same, the identity of who is administering targeting and enrollment is different. Previously, community management committees (CMCs) consisting of at least 10-12 members were used. Now, targeting is done using community teams, where a team consists of two people (one who can write easily, and one who is especially knowledgeable of the local context), and the total number of teams is determined by the expected number of poor beneficiaries in a village.
2. In the pilot, a beneficiary was a person (a child, an elderly person, etc). Now, the beneficiary is an entire household.
3. Previously, conditions applied to both children and the elderly. Now, they apply only to children. As we saw from the pilot evaluation, health visit conditions had little impact on the health of the elderly, but they improved the health of children. Imposing health conditions on only children therefore focuses on individuals who are most likely to benefit from more clinic visits, and whose human capital will be most positively affected.
4. Previously, education conditions were checked every 2 months. Now it is monthly. Given that the pilot evaluation showed minimal impacts on student absenteeism, more frequent and

vigilant monitoring and oversight may help deliver better compliance and attendance results. Also, health conditions for children aged 2-5 are now monitored half yearly rather than 6 times per year.

5. The benefit structure changed somewhat. It now consists of both a fixed and a variable component, rather than only a variable component. The fixed component is a “basic transfer” that is a fixed amount provided to guaranteed to all of those enrollment. Then there is also a variable component. It is only available if there is compliance among children related to health and education.
6. While previously, only students up to age 15 were monitored and provided support conditional on attendance at school, now support has been extended up to secondary education level (junior and senior secondary school). Support will even be given to individuals over age 18, as long as they are enrolled in secondary school. This was a decision we made after learning about the benefits of the CCT for helping children transition to secondary school — especially girls.

In addition to these design changes to the CCT based on the pilot evaluation, our questionnaires are being used in conjunction with the evaluation of the scale-up.

Community management committees receive a small allowance to cover the costs of travel and lodging, as they travel to retrieve and distribute funds and monitor conditions. However, they do not receive any wages or fees. At the outset of the program, these costs were estimated to come to approximately US\$5 per committee member per month, just under 5 percent of the total value of the transfers. Beyond the transfers themselves, other recurring costs included local bank transaction costs, equal to about 1 percent of the total value of the transfers. One-time or two-time costs included training on the program provided to all beneficiaries (about 12 percent of the total value of the transfers), consultations on the design and implementation with stakeholders at a number of levels, and printed materials (for training and for monitoring conditions). The estimated cost of the process of targeting beneficiaries came to less than 1 percent of the total value of the transfers.

References

- (2012). Tanzania Productive Social Safety Net Project. Project Appraisal Document. Report Number 67116. <http://documents.worldbank.org/curated/en/2012/03/15927283/tanzania-productive-social-safety-net-project>.
- 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.
- Aker, J. C. (2013). Cash or coupons? testing the impacts of cash versus vouchers in the Democratic Republic of Congo. *Center for Global Development Working Paper*.
- Akresh, R., D. De Walque, and H. Kazianga (2013). Cash transfers and child schooling: Evidence from a randomized evaluation of the role of conditionality. *World Bank Policy Research Working Paper* (6340).
- Akresh, R., D. de Walque, and H. Kazianga (2014, May). *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>.
- Arcand, J.-L. and L. Bassole (2007). Essential heterogeneity in the impact of community driven development. *processed, CERDI-CNRS, Université d'Auvergne, November*.
- Babbal, B. (2012). Evaluating equity in the provision of primary health care in Tanzania. <https://ir.library.oregonstate.edu/xmlui/handle/1957/32933>.
- Baird, S., E. Chirwa, C. McIntosh, and B. Ozler (2015). What happens once the intervention ends. *The 5-Year Impacts of a Cash Transfer Program in Malawi*. unpublished manuscript.
- 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>.
- Behrman, J., J. Hoddinott, et al. (2000). An evaluation of the impact of progresá on pre-school child height. *International Food Policy Research Institute, Washington, DC*.
- Behrman, J. R., P. Sengupta, and P. Todd (2005). Progressing through progresá: An impact assessment of a school subsidy experiment in rural Mexico. *Economic development and cultural change* 54(1), 237–275.
- 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.
- 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).
- Chase, R. and M. Woolcock (2005). Social capital and the micro-institutional foundations of cdd approaches in east asia: evidence, theory, and policy implications. In *Arusha conference, "New Frontiers of Social Policy," December*, pp. 12–15.
- Costa, A. B. M., M. Gyoeri, and F. V. Soares (2016). Community of practice on cash transfers in africa. http://www.ipc-undp.org/put/eng/PRB_CoP01_Scaling_up_cash_transfer_programmes_good_practices_and_lessons.pdf
- Davis, B., S. Daidone, J. Dewbre, and K. Covarrubias (2015). The impacts of the child grants programme in lesotho.
- 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).
- 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.
- Edmonds, E. V. (2006). Child labor and schooling responses to anticipated income in south africa. *Journal of development Economics* 81(2), 386–414.
- 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., M. Park, R. O. Darko, I. Osei-Akoto, B. Davis, and S. Daidone (2013). Livelihood empowerment against poverty program impact evaluation. *Chapel Hill: University of North Carolina, Carolina Population Center*. Available at: http://www.unicef.org/ghana/gh_resources_LEAP_Quant_impact_evaluation_FINAL_OCT_2013.pdf
- 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 (2013). Household response to income changes: Evidence from an unconditional cash transfer program in Kenya. *Massachusetts Institute of Technology*.
- Independent Evaluation Group (2011). *Social safety nets: An evaluation of world bank support, 2000–2010*. Washington, DC: Independent Evaluation Group, the World Bank Group.
- 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.
- Macours, K., N. Schady, and R. Vakis (2008). Cash transfers, behavioral changes, and the cognitive development of young children: Evidence from a randomized experiment. *Policy Research Working Paper 4759*.
- Manley, J., S. Gitter, and V. Slavchevska (2013). How effective are cash transfers at improving nutritional status? *World Development* 48, 133–155.
- Mansuri, G. and V. Rao (2013). *Localizing development: does participation work?* World Bank Publications.
- Marriott, A. (2011). Does health insurance work in Tanzania? <http://www.globalhealthcheck.org/?p=388>, accessed December 15, 2014.
- 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.
- Natali, L., S. Handa, D. Seidenfeld, G. Tembo, et al. (2015). The impact of Zambia's unconditional child grant on schooling and work: Results from a large-scale social experiment. Technical report.
- 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. *Forthcoming, Journal of Health Economics*.
- Rawlings, L. B. and G. M. Rubio (2005). Evaluating the impact of conditional cash transfer programs. *The World Bank Research Observer* 20(1), 29–55.
- 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). xtvreg2: 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>.
- Uwezo (2014). Are our children learning? literacy and numeracy across east Africa: 2013.
- Ward, P., A. Hurrell, A. Visram, N. Riemenschneider, L. Pellerano, C. O'Brien, I. MacAuslan, and J. Willis (2010). Cash transfer programme for orphans and vulnerable children (ct-ovc), Kenya.

World Bank (2013). Service delivery indicators at a glance.

World Bank (2014). Big results now in education program: Program appraisal document (report no. 84545-tz).

World Bank (2016). World development indicators.

World Health Organization (2014). United Republic of Tanzania: Health Profile. *Africa's Pulse*. <http://www.who.int/gho/countries/tza.pdf?ua=1>, accessed December 15, 2014.

Zeza, A., B. de la Brière, and B. Davis (2010). The impact of social cash transfers on household economic decision making and development in eastern and southern africa. In *African Economic Conference*, pp. 27–29.

Appendix A: Pre-analysis Plan

TW1.1001-Evaluating the Effectiveness of a community-managed Conditional Cash Transfer Program in Tanzania

Katrina Kosec

David Evans

Stephanie Hausladen

Pre-analysis Plan — September 2012

As our 3ie application described, the main evaluation questions our study proposed to address are the following (listed in italics). Our associated pre-analysis plan follows each italicized item.

1. *Is the community-administered CCT model an efficient method to achieve health, education, and consumption gains? How does its cost effectiveness compare to that of more centralized CCT programs in similar contexts?*

We will compare our evaluation with other cash transfer programs through an analysis of how cost-effective other evaluated cash transfer programs in sub-Saharan Africa have been to date. We will summarize the findings of this research effort in a summary table, to be submitted with our next progress report. While it is always a challenge to compare the cost effectiveness of programs with different objectives and with evaluations done at different points in time, this will lend some insights (and provide some orders of magnitude) that can help us assess this community-managed CCTs relative effectiveness.

2. *What are the specific impacts of this program on the following outcomes?*

- Household outcomes (health, education, time-use, consumption, transfers, savings, asset accumulation, decision-making, intra-household allocation, attitudes, and trust)
- Welfare of particular groups (examine heterogenous impacts)
- Community dynamics (member relationships, social capital, traditional solidarity systems, conflicts)

While villages were randomly assigned to treatment, some imbalances between treatment and control households at baseline highlight the importance of an empirical strategy that takes initial conditions into account. We thus use a difference-in-differences estimator with the following empirical specification:

$$\text{Outcome} = \beta_1 + \beta_2(\text{After}) + \beta_3(\text{Treatment} \times \text{After}) + \Gamma(\text{Unit_of_Obs_Fixed_Effects}) + E \quad (2)$$

Outcome is a household or individual-level potential impact of treatment, Treatment indicates receipt of treatment (it takes the same value in both periods), and After indicates that the observation comes from a post-baseline survey. *Unit_of_Obs_Fixed_Effects* are dummies for the household (in the case of a household-level regression) or the individual (in the case of an individual-level regression). They ensure that we take into account all time-invariant household or individual characteristics (not just those we can measure), in case treatment is somehow correlated with them. β_3 is then the effect of having received treatment.

We will use two definitions of treatment to yield two estimates of β_3 . First, we will consider treatment to mean living in a treatment village (where all of those surveyed were intended to be beneficiaries of the program). This provides “intent to treat” (ITT) estimates of β_3 . As

our baseline survey occurred before randomization and before rules for treatment had been finalized, some people we surveyed in treatment villages were ultimately not treated. The ITT analysis bases inference off random assignment to treatment, not whether treatment was actually received (which may be for non-random reasons). Second, we will consider treatment to mean receiving transfers. This provides “effects of treatment on the treated” (TOT) estimates of β_3 . The TOT analysis circumvents the problem of some people failing to receive treatment for non-random reasons by instrumenting for receipt of treatment with (random) assignment to treatment.

In addition to fixed effects analysis, the team notes that applying fixed-effect models also implies an assumption of parallel trends for the treatment and the control group. This assumption can usually be made without danger when randomization is successful. However, under the current study’s circumstances (with several variables that were not balanced at baseline), it remains questionable whether parallel trends for the treatment and the control group can be assumed. The team will complement the fixed-effect analysis with another technique, in which the first difference of the outcome variable is regressed on both the treatment assignment and the variables that are not balanced at baseline.

Specific outcomes we will examine include the following, motivated by the theory of change detailed in our application:

- Health-seeking behavior
- Health, nutrition, and the ability to perform daily tasks
- Consumption quantities and values
- Amount and composition of expenditures
- School attendance and education outcomes
- Types of activities children perform
- Health- and education-promoting assets owned by children
- Household savings and credit decisions
- Trust (of people overall, people in their community, and community leaders) and satisfaction with community services
- Inter-household transfers
- Asset purchases (especially land and livestock)

We measure health seeking behavior through questions on the household questionnaire asking about the most important kind of health provider visited for the main health problem, if any medication was taken, and how treatment was financed. We ask how many times each individual beneficiary visited a health center in the past year, when they last visited, how far away this facility is, if they would visit this facility in the absence of the CCT program, why they visited the health facility, the cost of treatment, and if they had to return home without being attended. We have also added questions on participation in the Community Health Fund, a type of household health insurance. In addition, we also ask health providers at each health center or dispensary to answer questions on how many patients their center has seen in the past month.

3. *What are the major factors in program implementation that contribute to the success and efficiency of this model? What are potential bottlenecks to effective implementation? How does community capacity affect program success?*

We want to examine three major types of heterogeneous impacts, as well as some additional analyses. We feel that knowledge about each of these will generate valuable information about when and where a community-managed CCT program could have the greatest impact. These include whether a CCT program is more effective in places:

- That have experienced severe adverse economic and climate shocks
- With strong underlying governance institutions and good public service delivery
- With more unequal distributions of income and with more poverty

First, we can imagine transfers being more effective in the presence of severe, adverse economic shocks like droughts, floods, and other exogenous sources of crop failure. Such shocks can erode existing savings, lead to the liquidation of assets, and even threaten households' ability to rely on informal insurance arrangements if shocks are felt community-wide. This may make education and health investments less affordable. For shock-affected households, a little extra income may be especially marginal to whether basic health, education, and nutrition needs can be met.

Second, we can imagine the quality of governance and the quality of public service delivery having a dramatic impact on what parents can achieve with a given amount of transfer money. If these institutions are effective, then people may receive more money, at more regular intervals, and make longer-term investments since they are more assured of a steady stream of income. Further, parents may perceive education and health investments to be higher-value (e.g., because children learn more at school, or receive supplementary nutrition at school, or get better diagnoses and treatments at health facilities, etc.)

Third, we can imagine transfers being more effective for poor people (for whom a little extra money is more likely to influence whether they make basic education and health investments). Also, we can imagine them being more effective for those that live in communities with greater income inequality, as such individuals may have fewer social networks and neighbors at the same economic level with whom they can form mutually-beneficial informal insurance arrangements; on the other hand, this may not be the case. The team will be precise with what our variables measure and do not measure. Instead of simply "inequality," the team will use variables that more directly link with our theory of change, such as what share of the population is beneficiaries and what is the level of trust people have in other community members.

Furthermore, in addition to the three types of heterogeneous effects already listed, we will also investigate differential effects by parental education levels, distance to health facilities and schools, satisfaction with health and schooling facilities at baseline, enrollment status at the baseline, gender, and characteristics of the head, such as education levels, gender, age, and occupation. We will also look at effects of the availability of water, which was one of the main problems cited by villagers during our field visits and can prevent children from attending school because they must fetch water instead. Additionally, we will attempt to construct measures of community trust and village governance, and differentiate results at these levels as well.

4. What are the spillover effects of the program? How do they compare to direct effects?

We will study spillovers by econometrically analyzing how treatment affects informal transfers of various types (cash, food, other in-kind, or labor) between households. We wish to assess what share of transfers is kept by beneficiary households, and what share seems to be paid out to non-beneficiary households.

5. *Can compliance with program conditions be linked to the welfare of children and the elderly? Or do the effects come through other mechanisms? (analysis of theory of change)*

A thorough analysis, such as that described in points 2 and 3 (above), will offer initial insights into which mechanisms are generating program benefits. It will make clear what are the effects of treatment and under what conditions. However, as we described in our theory of change, we plan to do additional analysis inspired by Macours, Schady, and Vakis (2008) to assess to what extent benefits are due to program-induced behavioral changes vs. income effects.

6. *How much do these transfers protect households from shocks vs. assist them in moving out of poverty?*

Our analysis of the heterogenous impacts of treatment across households that did and did not experience severe, adverse economic shocks will help us understand if the program has effects aside from those accruing to households experiencing shocks.

7. *Do community score cards improve the effectiveness of these transfers?*

Twenty randomly-selected treatment villages received community score-cards and 20 did not. We will provide some analysis of whether, at endline, outcomes were better among treatment communities that did the scorecard exercise.

8. *Does transmitting payments through mobile phone banking reduce leakage?*

This is not something we have been able to explore yet, though we have had discussions with the government. We will continue to discuss the feasibility of introducing mobile phone banking with the government, and will keep 3ie aware of any developments.

Appendix B: Stata analysis code

In this report, results are reported from three different models. Below we provide the Stata (Stata 14.0) code used to estimate the effects. Here the variables are defined:

Y = the outcome being tested

receivedtreat2 = 1 if individual lived in treatment household during midline survey; 0 otherwise

receivedtreat3 = 1 if individual lived in treatment household during endline survey; 0 otherwise

assignedtotreat2 = 1 if individual lived in treatment village during midline survey; 0 otherwise

assignedtotreat3 = 1 if individual lived in treatment village during endline survey; 0 otherwise

D2 = 1 if the midline; 0 otherwise

D3 = 1 if the endline; 0 otherwise

V = a variable that uniquely identifies each of the 80 villages

8.0.1 Individual-level outcome with person fixed effects

Estimate the effect of assignment to a treatment village:

```
use "individual_dataset", clear
iis person_id_var
xtreg Y assignedtotreat2 assignedtotreat3 D2 D3 , fe cluster(V)
```

Estimate the effect of treatment on the treated:

```
use "individual_dataset", clear
iis person_id_var
xtivreg216 Y (receivedtreat2 receivedtreat3=assignedtotreat2 assignedtotreat3) D2 D3, fe cluster(V)
```

8.0.2 Household-level outcome with household fixed effects

Estimate the effect of assignment to a treatment village:

```
use "household_dataset", clear
iis household_id_var
xtreg Y assignedtotreat2 assignedtotreat3 D2 D3 , fe cluster(V)
```

Estimate the effect of treatment on the treated:

```
use "household_dataset", clear
iis household_id_var
xtivreg2 Y (receivedtreat2 receivedtreat3=assignedtotreat2 assignedtotreat3) D2 D3 , fe cluster(V)
```

¹⁶Schaffer (2010)

8.0.3 Household-level outcome with control set

One outcome in this report (household participation in the Community Health Fund) is only available in the endline survey. For that reason, it is not possible to use household fixed effects to estimate the effect of treatment. Instead we use a set of baseline controls. The controls include 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 an asset index constructed from the first principal component from a PCA using information on ownership of 13 household assets.

Estimate the effect of assignment to a treatment village:

```
use "household_dataset", clear  
reg Y assignedtotreat3 'controls' , cluster(V)
```

Estimate the effect of treatment on the treated:

```
use "household_dataset", clear  
ivregress 2sls Y (receivedtreat3= assignedtotreat3 ) 'controls' , cluster(V)
```

Appendix C: Questionnaires