# In search of the holy grail: Can unconditional cash transfers graduate households out of poverty in Zambia?

Sudhanshu Handa (Co-PI)
University of North Carolina at Chapel Hill

Gelson Tembo (Co-PI)
Palm Associates and University of Zambia

Luisa Natali
UNICEF Office of Research-Innocenti

Gustavo Angeles
University of North Carolina at Chapel Hill

Gean Spektor
University of North Carolina at Chapel Hill

**Grantee Final Report**

**Accepted by 3ie: August 2019**

## Note to readers

This impact evaluation has been submitted in partial fulfilment of the requirements of grant DPW1.1042 issued under Development Priorities Window 1. This version is being published online as it was received. A copy-edited and formatted version will be available in the 3ie Impact Evaluation Report Series in the near future.

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, Sudhanshu Handa at mshanda@email.unc.edu.

The 3ie technical quality assurance team comprises Neeta Goel, Deeksha Ahuja, Sayak Khatua, Marie Gaarder, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder.

Funding for this impact evaluation was provided by UK aid through the Department for International Development. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Handa, S, Tembo, G, Natali, L, Angeles, G, and Spektor, G, 2019. In search of the Holy Grail: can unconditional cash transfers graduate households out of poverty in Zambia, 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie)

## Acknowledgements

This research is jointly funded by the International Initiative for Impact Evaluation (3IE), UNICEF-Zambia, and UNICEF Office of Research through support from Sida. The evaluation was conducted by the University of North Carolina at Chapel Hill, USA and Palm Associates, Lusaka, Zambia. Luisa Natali is employed by UNICEF Office of Research. Thanks to Averi Chakrabarti and Valerie Lundeen for excellent research assistance, and to the 3IE referees for useful comments. Sudhanshu Handa and Gelson Tembo are the corresponding authors for this report. Manzunzu Zulu (manzunzo@yahoo.co.uk) is the focal point of the study for the Ministry of Community Development & Social Services, the implementing agency.

## Summary

The study assesses whether the large, positive impacts of the Child Grant Program (CGP) were sustained after households exited the programme. The CGP was an unconditional cash transfer that provided approximately US$12 to households with a child under age five in three rural districts of Zambia. The study exploits the fact that the Government of Zambia reformed its grants programmes in 2015 such that many households who had been in the CGP since 2011 were no longer eligible for the new programme, and effectively were removed from the rolls. We compare these households to those in an experimental control group who were never enrolled in the CGP. The analysis is complicated by the fact that CGP households were gradually removed from the programme so that some received benefits longer than others. Even accounting for this variation in programme exposure, we find that the consumption and food security of the original CGP households declined after the programme ended, and that there was no longer any difference between them and the original control group. These findings are supported by indicators of subjective well-being, which also drop for the treatment group and are equal to those among the original control group.

There is some nuance to these results when looking at productive activity and assets. The level of these indicators stays the same or even increases among the original treatment group, but increases even more in the control group, again leading to convergence, but convergence that is driven by ‘catch-up’ rather than ‘fade-out’. The pattern of results on assets and productive activity suggests that the original treatment households are economically stronger, more resilient, and perhaps more likely to withstand shocks to consumption, though we cannot equivocally assert that from the results of this analysis.

Households in this study are ultra-poor, with mean consumption of US$0.30 per person per day, some of the poorest households in the world. Infrastructure and environmental conditions are likewise quite harsh. The pattern of results we find, that large programme effects mostly fade-out quickly, suggest that in this environment, and with households at the edge of survival, cash alone is unlikely to lead to wholesale graduation out of poverty.


> Grail
>
> An object or goal that is sought after for its great significance. — Meriam-Webster Dictionary
>
> A thing which is eagerly pursued or sought after. — Oxford English Dictionary

# Contents

## List of figures and tables

## Abbreviations and Acronyms

| | |
|---|---|
| CGP | Child Grant Program |
| CWAC | village clusters |
| DD | difference-in-difference |
| GDP | gross domestic product |
| GoZ | government of Zambia |
| HSCT | Harmonized Social Cash Transfer |
| LCMS | Living Conditions Monitoring Survey |
| MCDSS | Ministry of Community Development & Social Services |
| NGO | non-governmental organization |
| PMT | proxy means test |
| RCT | randomized controlled trial |
| SCT | social cash transfer |
| SD | standard deviation |
| SSA | sub-Saharan Africa |
| UCT | unconditional cash transfer |
| ZDHS | Zambia Demographic and Health Survey |
| ZMW | Zambian kwacha |

# 1. Introduction

A recent review by the World Bank (2015) estimates that around 150 countries in the developing world have cash assistance programmes and that approximately 800 million people are reached by some type of cash transfer programme. Significant expansion of cash transfer programmes has recently occurred in sub-Saharan Africa (SSA), with a doubling of development oriented (i.e. non-humanitarian) programmes from 20 to 41 between 2010 and 2015, reaching an estimated 8-10 million households or 50 million individuals (World Bank 2015; Garcia and Moore 2012). These trends have raised questions about the long-term implications of cash transfer programmes and their affordability in low-income settings. There is increasing interest in finding out whether such programmes can 'pay for themselves' by generating inclusive growth which raises GDP and national revenue. In other words, are poverty-focused cash primarily for protective purposes, necessitating other, integrated programmes focused explicitly on livelihoods for those who are capable of graduating out of poverty? Or can these programmes also generate productive effects that can lead to longer-term graduation from poverty?

Recent discussions around interventions to achieve sustained poverty reduction in developing countries have centred around two broad approaches. The integrated livelihoods approach, pioneered by BRAC, is an intensive big push, which provides a range of services to the ultra-poor, including cash, assets, and livelihoods training. This approach is costly, but a recent paper reports on the results from six RCTs of the BRAC model suggest that the cost-benefit ratio can be quite large, and could be made even larger if some of the more costlier components of the programme that seem to have less of an impact can be removed (Banerjee et al 2015). The paper shows sustained impacts on consumption one year after the intervention ended, suggesting that this model may lead to sustained graduation out of poverty, and thus potentially represents the Holy Grail. The key challenge with the BRAC model however is that it is complex, and currently only ever implemented by non-governmental organisations, leading to serious questions about scale-up potential, and whether it can ever be part of a national social protection system.

At the other end of the spectrum are proponents of unconditional cash transfers to the ultra-poor. A recent article in Foreign Affairs (Blattman et al 2014) argue that unconditional cash should be the new benchmark in foreign aid, that very few interventions can beat the cost-effectiveness of providing the ultra-poor with plain cash, which allows them to spend money in the way that best allows them to satisfy their priorities. Results from rigorous evaluations of national UCT programmes in Malawi (University of North Carolina, 2016), Zambia (Handa et al. 2016a) and Zimbabwe (University of North Carolina 2018) do find impacts across a range of protective and economic domains. And Give Directly, an NGO distributing unconditional cash grants in three lump-sum payments also report similar protective and productive impacts in Kenya (Haushofer & Shapiro 2016). However, it is not known whether these effects are long lasting and in do fact lead to graduation out of poverty.

There is a small but growing literature that assesses the long-term poverty effects of cash or near-cash interventions. Stoeffler et al. (2016) show that an 18 month unconditional transfer pilot in rural Niger had sustained impacts on assets and productive

activities *18-months* after transfers ended. Evidence on the post-intervention effects of a one-year scholarship combined with a conditional cash transfer on schooling and child labour outcomes in Nepal is less positive, indicating no permanent impacts 17 months after the final disbursements (Edmonds and Shrestha 2014). Baird et al (2018) investigate the durability of impacts from two-year conditional and unconditional transfers targeted to adolescent girls and young women in Malawi, returning more than two years after the termination of the programme. Most impacts dissipated over time in the unconditional cash transfer treatment arm: still, authors found an impact on the height for age scores of children born during the programme. The conditional arm had some long-lasting effects, including on educational attainment and total number of births—interestingly, only for girls who were out-of-school at baseline. That study does not look at economic or productive outcomes however. Haushofer and Shapiro (2018) find a no strings attached one-time lump cash payment by GiveDirectly had sustained impacts on assets but not on other outcomes three years after the programme began[1]. Therefore, although we have some evidence from post-intervention studies, it is mixed and with few examples across geographic regions, programme types and outcome domains.

We contribute to the evidence on the graduation potential of cash transfer programmes by taking advantage of a unique reform in Zambia's cash transfer programme where a significant number of households were no longer eligible for the Child Grant Program (CGP) after five years. A large randomised control trial (RCT) reported significant protective and productive effects of this unconditional cash transfer (UCT) on households, with estimated multiplier effects in the range of 1.5 (Handa et al 2018). Do these effects persist after the programme ended? What has happened to consumption and asset accumulation among households that are no longer on the programme? Answering these questions will help us understand whether UCTs represent a viable option for governments to address current and future poverty or whether these programmes simply address short-term protection without representing a real pathway out of poverty.

This report is structured as follows. Section 2 describes the original intervention, the theory of change and the research questions. Section 3 presents the research design, core identification strategy and empirical specifications. Section 4 presents the main findings, and Section 5 discusses the results, policy implications and concludes.

# 2. Intervention, Theory of Change and Research Questions

## 2.1 Intervention

In 2010 the Government of Zambia's Ministry of Community Development and Social Services (MCDSS) decided to implement two cash transfer demonstrations with slightly different targeting criteria to inform future scale-up. Both programmes underwent rigorous impact evaluations using experimental designs (Handa et al 2018). The

[1] This is technically different than cash transfers which are usually defined as regular and predictable monetary transfers to poor and/or vulnerable populations. There are then also studies on programmes that provide cash (usually a one-time lump sum) conditional on business start-up. These are conceptually different programs and we do not cover them here (see for instance Blattman et al 2018).

programme we study in this report is the Child Grant Program (CGP), which targeted any household with a child under five years old[2] in the three districts of Shangombo, Kalabo and Kaputa. Eligible households originally received ZMW 55 a month (equivalent to U.S. $12) irrespective of household size, an amount deemed sufficient to purchase one meal a day for everyone in the household for one month. Though not explicitly poverty targeted, these districts are extremely poor such that 90 percent of CGP recipients were below the national extreme poverty line, and median consumption prior to programme start-up was 30 US cents per person per day. The transfer amount represented approximately 25 percent of baseline consumption, slightly higher than similar programs in Malawi and Zimbabwe.

In 2015 the GoZ consolidated these two and several other existing cash transfer programmes into one national harmonised 'inclusive' model and began a rapid scale-up. New districts were targeted, and beneficiaries in demonstration districts were slowly re-targeted as well. The 2016 model targeted incapacitated labour-constrained households, specifically households with an elderly member and/or a member with a severe disability. A welfare criterion was applied in the second stage of the screening process in order to filter out the better-off households[3]. While the welfare criterion is implemented using a proxy means test (PMT), it differs significantly from the typical application of the PMT. Rather than using the PMT to try and identify the ultra-poor and include them in the programme, the PMT is used in Zambia to identify the ultra-rich and exclude them from the programme. This approach essentially eliminates exclusion error based on relative poverty, a sharp criticism of the PMT approach (Brown, Ravallion & van de Walle, 2016). As of November 2017 the transfer was set at a flat unconditional ZMW 90 per month, paid in cash every two months. Households with a disabled member received twice this amount. By the end of 2017, the harmonised programme had reached approximately 550,000 households representing 3m individuals or 18 percent of the population.

In the 2017-18 budget speech the Minister of Finance pledged additional funding for the Harmonized Social Cash Transfer (HSCT) to reach 700,000 households and so new beneficiaries are being added to the rolls. The 2017-2018 budget allocation is ZMW 721 million, which represents approximately 85 percent of the programme's overall budget, the remaining 15 percent coming from foreign aid, primarily in the form of technical support, monitoring and evaluation.

2 All households with a child under the age of five were eligible for the Child Grant Program. However, for the evaluation, only households with a child under three years old were considered, the rationale being that these households would have then been covered by the program for at least two years.

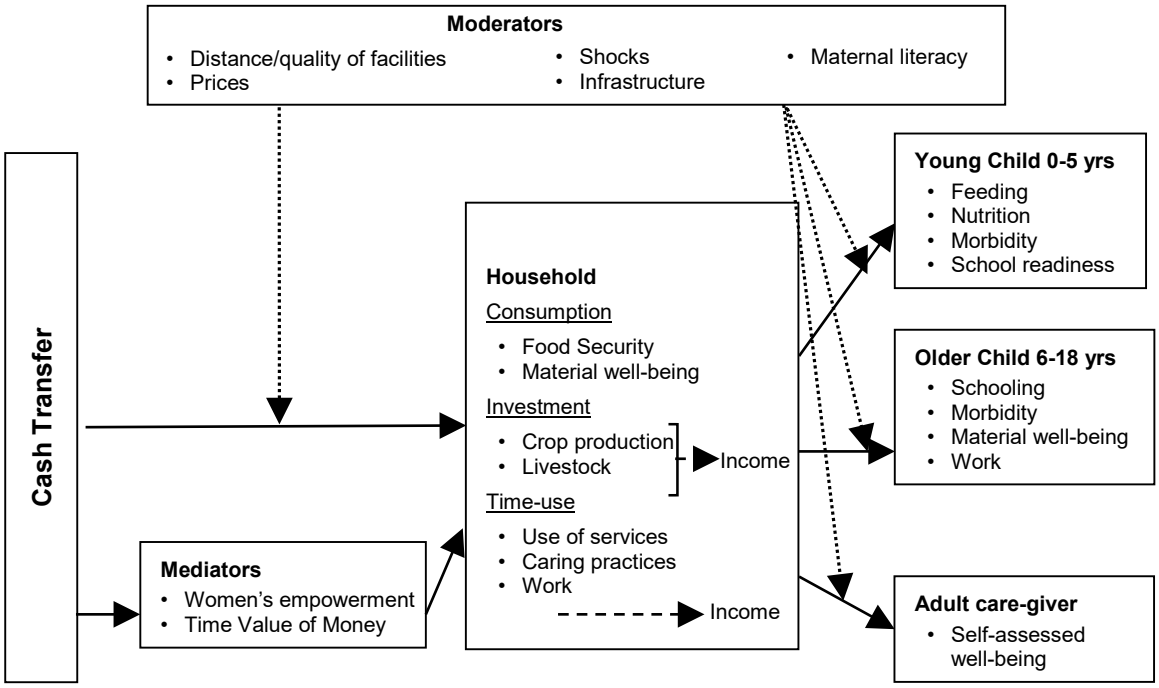
3 Since 2016, the two-stage targeting model has been modified based on observations from the initial implementation. For example, information collected in the first stage is now more comprehensive, aiming to list all households hosting persons with severe disability, households with elderly members 65 years or older, child-headed households, female-headed households with three or more children under the age of 19 years, and/or households with chronically ill patients on palliative care. As in 2016, a welfare criterion is used in the second stage to screen out all households that were relatively well-off regardless of their eligibility standing from stage 1.

## 2.2 Theory of Change

This primary objective of the current study is to assess the effects of the CGP after recipients have left the programme study. It is not a study about the impact of the new harmonised programme, so below we lay out the theory of change of the CGP.

The CGP provided an unconditional cash transfer to households with a child under age 5. The theory of change or conceptual framework for the impact evaluation was developed with stakeholders and programme implementers at several workshops as part of the activities of the original impact evaluation of the programme in 2010. These discussions informed the survey instruments and main indicators to be collected. The final conceptual framework agreed upon for the evaluation is shown in Figure 1 and is read from left to right. We expected a direct effect of the cash transfer on household consumption (food security, material well-being), on the use of services, and possibly even on productive activity after some time. Based on sociological and economic theories of human behaviour we posited that the impact of the cash might work through several mechanisms (mediators), including a woman's bargaining power within the household (because the woman receives the cash directly) and the degree to which the woman is forward looking (time value of money). Similarly, stakeholders and programme implementers agreed that the impact of the cash transfer may be weaker or stronger depending on local conditions in the community. These moderators include access to markets and other services, prices, and shocks. Moderating effects are shown with lines that intersect with the horizontal lines to indicate that they can influence the strength of the direct effect.

**Figure 1: Conceptual framework for impact evaluation of Child Grant Program**




Note: Source is CGP Baseline Evaluation Report available at https://transfer.cpc.unc.edu/?page_id=1262

In the figure we list some of the key indicators along the causal chain that were analysed in the evaluation of the CGP. These were consistent with the results framework of the project and were all measured using established items in existing national sample surveys such as the Living Conditions Monitoring Survey (LCMS) and the Zambia Demographic and Health Survey (ZDHS). A key strategic decision made by stakeholders and programme implementers was to administer the full consumption module from the LCMS in the evaluation survey, which covered close to 250 individual items. This was done because consumption expenditure was viewed as the single most important welfare measure for the programme, and because ultimately all household decisions in the long-run would show up in consumption. Though extremely time-consuming to administer, this enabled the study team to track in detail the pattern of consumption of households and understand how the cash filtered through the household economy.

There were several key assumptions underlying the conceptual framework: the full amount of cash would be delivered on-time as per programme operational rules; that markets existed and were sufficiently responsive to increases in demand so that households could purchase food and other basic needs, and prices would not rise due to supply constraints thus eroding value of transfer in local markets; purchasing power of cash would not be eroded due to general (country-wide) inflation. The study team tracked these assumptions through the course of the original evaluation. Data from the programme management information system confirmed regular bimonthly payments through the life of the original study between 2010 and 2014. A detailed operations manual administered to beneficiaries did not uncover any systematic concerns of bribery or withholding of cash. The fact that the amount was flat made it easy for recipients to know how much they were entitled to receive, and to confirm the amount. A detailed community price questionnaire captured prices in treatment and control communities and revealed no price inflation in treatment clusters. Finally, the value of the transfer was raised through-out the initial four year study period, initially starting at ZMW 55 per month and rising to ZMW 70 by the end of 2014 to keep up with general inflation in the country. As mentioned earlier, the transfer value has since been raised to ZMW 90.

## 2.3 Research Questions:

This study is motivated by the finding from the original evaluation of the CGP that the programme had generated a significant multiplier effect on recipient households in the range of 1.5 by 2013, three years after the transfers started. In other words, each kwacha (ZMW) transferred was converted into an additional 0.5 ZMW, primarily through non-farm enterprise and increased agricultural production (Handa et al 2018). The harmonization of the cash transfer programmes in Zambia meant that many of the original CGP beneficiaries would no longer be eligible for the new programme, providing a unique natural experiment to assess whether the original impacts were sustained after the transfer was removed. The specific research questions are:

1) What happens to the consumption and economic and financial position of households who originally received the CGP but do not receive the HSCT? Do the original impacts fade-out or are they sustained?

2) A related question is that of catch-up. Do households in the original control group who are eligible for the HSCT show levels of consumption and economic and financial

positions that are the same as the households that received the CGP and also were eligible for the HSCT? These latter households in principle continued to receive transfers for almost seven years, although as we show below, some of these households had exited the CGP for some time because their focal child had aged out of the programme.

**Key Terms**

Catch-up: Both the T and C groups improve, but the improvement is greater in the C group, so they catch-up to the T group.

Fade-out: The initial advantage of the T group over the C group is eliminated due to a subsequent reduction in the T group.

Convergence: The T and C groups start at different levels but end up at the same level. This can occur because of fade-out or catch-up.

# 3. Evaluation Design, Sampling and Data

## 3.1 The Original Impact Evaluation

The initial impact evaluation of the CGP that this study builds on was a multisite RCT conducted in the three programme districts of Kalabo, Shangombo and Kaputa selected based on their high poverty and child malnutrition rates. An inception meeting to discuss and agree upon evaluation design options was held in June 2010 at the Ministry of Community Development & Social Services (MCDSS), and included provincial and district social workers from the programme sites. Subsequently the ministry conducted the first step of the randomization process by randomly selecting 30 CWACs (village clusters) within each district, out of roughly 100 CWACs in each district, to enter into the study through a lottery held at the Ministry headquarters. This is the *random selection* of study sites.

After the 90 CWACs (30 from each district) were randomly selected for the study, targeting in each study CWAC was undertaken. Ministry staff (social welfare officers), CWAC members and staff from the district health services (nurses, community health workers) identified all eligible households with at least one child under 5 years old in the study communities. The identification process entailed house to house visits, coupled with public awareness campaigns. This process resulted in more than 100 eligible households in each CWAC on average. From this master list, 28 households were randomly sampled from each CWAC for inclusion in the study with four additional households per CWAC kept in reserve in case of refusals. Baseline data was then collected in October 2010 on 2519 households prior to the random assignment to intervention or control status; households were thus blinded at baseline.

After the baseline was completed, *random assignment* to study arms was conducted in public with local officials, Ministry staff, and community members present as witnesses. Within each district, CWACs were randomly ordered using the ‘random’ function in Microsoft Excel. The Permanent Secretary of the MCDSS then flipped a coin to determine whether the top half or the bottom half of the list would enter the programme first. The randomization was stratified by district, so control and treatment CWACs were

drawn from the same districts. The level of randomization is the CWAC, so treatment and control households do not live next to each other in the same village, thus reducing the possibility of spillover or other forms of contamination.

The original evaluation included four CGP follow-up waves at 24, 30, 36 and 48 months. All surveys were conducted in October-November to control for seasonality, except for the 30-month survey, which was purposely collected during the harvest season to assess consumption smoothing effects. Evaluation reports for each survey round are available on the Transfer Project website, along with all survey instruments and related documentation. Additional details on sampling and fieldwork can be found in the baseline report (American Institutes for Research, 2011). The multiplier results are contained in the 36-month report. A study flow-chart and timeline are provided in Appendix A/B that includes the 2017 survey (84-months post CGP baseline). A map of the study districts is shown in Appendix C. Note that between 2013 and 2014 Zambia went through a redistricting exercise. A portion of the original district of Kaputa became a new district (Nsama), and likewise a portion of the original district of Shangombo became a new district (Sioma) and a portion of Kalabo district became Sikongo. However, we maintain the original three districts as our stratification variable in the analysis below. During the redistricting, two CWACs originally assigned to Kalabo ended up in Shangombo.

## 3.2 The Current Study

We returned to the original CGP evaluation sample in October-December 2017 and administered the identical survey instruments (household, community) using the same field teams and protocols as in prior waves. Our original sample at baseline comprised 2519 households; our final analysis sample uses a balanced panel of households who appear at baseline, 2013 and the 2017 follow-up, a total of 2109 households. Our pre-analysis plan specified that we would replicate the indicators used by Handa et al (2018) except for child nutrition, which was not collected in 2017. That article presents impacts of the CGP across eight domains encompassing consumption, food security, assets (productive, domestic and livestock), ‘finance and debt’, ‘incomes and revenues’, relative or subjective poverty, children’s material needs and schooling. To address the issue of multiple testing, we follow the same approach of constructing lead indicators or indexes for each domain and estimating impacts on those indexes or lead indicators. The precise variables (and definitions) used as the lead indicator or to construct the domain index are shown in Appendix D and are identical to Handa et al (2018). All indexes or lead indicators are standardised using the control group mean and standard deviation at each wave so that effect sizes can be readily compared across specifications.

## 3.3 Identifying Treatment Status

A major challenge in this study is accurately identifying the eligibility status of each sample household for the new Harmonized Cash Transfer programme (HSCT). In addition, the original CGP eligibility criteria included a provision that when the focal child turned six, they would no longer be eligible for the cash transfer. However, this ‘graduation’ policy was inconsistently implemented across the three districts and CWACs. Variation in graduation dates has important implications for our study because some treated households would have received cash payments under the CGP for much longer than others, which would potentially influence fade-out. And therefore, some

households would have stopped receiving the CGP far earlier, remaining uncovered for far longer and therefore increasing likelihood of fade out.

Table 1 summarizes the CGP end dates and HSCT retargeting and payment start dates in the study sites. The MCDSS purposely chose to retarget in control CWACs first and continued to make payments under the CGP and graduate households in the original treatment CWACs. The implications of this approach are quite important. For example, in the original districts of Shangombo and Kalabo, the retargeting under the HSCT in treatment CWACs only took place in October 2017 as we were going to the field, and no payments to HSCT-eligible households had ever been made in these CWACs, though they continued to receive CGP payments through the first few months of 2017. On the other hand, retargeting in control CWACs occurred over a year earlier and HSCT-eligible households in these districts had been receiving cash payments for over a year by the time we went to the field. *A comparison of the HSCT-eligible households across control and treatment CWACs in these districts would be comparing households who had been in the new programme for one year versus households who had not received payments for at least six months, and probably much longer if they had graduated earlier. This is a key empirical challenge in the study.*

**Table 1: CGP End Dates and HSCT Retargeting Dates in the Study Sites**

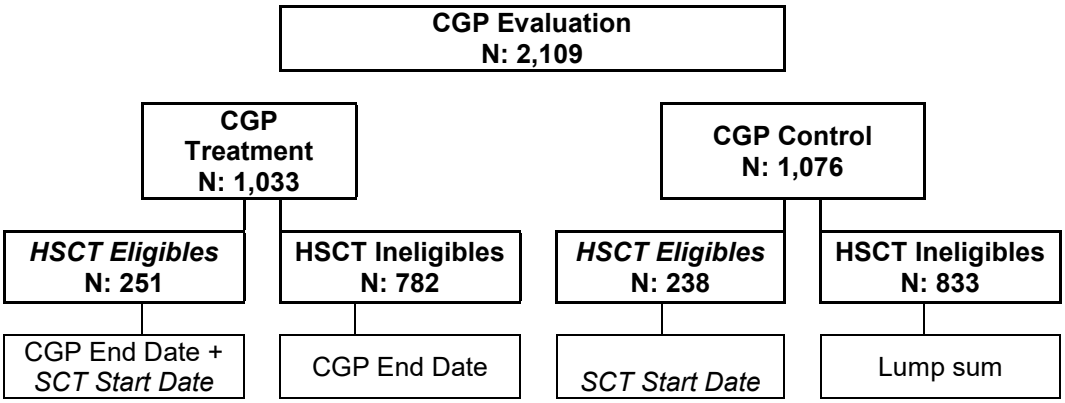
| District | # CWACs | CGP End Date | HSCT Start Date | |
|---|---|---|---|---|
| | | | Control CWACs | Treatment CWACs |
| Kaputa | 15 | February 2015 | June 2016 | May 2017 |
| Nsama (orig. Kaputa) | 15 | February 2015 | June 2016 | May 2017 |
| Shang'ombo | 17 | February 2017 | May 2016 | *Expected* January 2018 |
| Sioma (orig. Shang'ombo) | 15 | Q1 2017 | June 2016 | *Expected* January 2018 |
| Kalabo | 23 | February 2017 | May 2016 | *Expected* January 2018 |
| Sikongo (orig. Kalabo) | 5 | May 2017 | May 2016 | *Expected* January 2018 |

For both CGP graduation dates and household qualification for the HSCT, we compared our household sample with programme rosters at the district headquarters and hand match records based on CWAC and beneficiary name or name of household head. This process, though cumbersome and time-consuming, proved effective at identifying households in our sample who were eligible for the HSCT—there were just five households with missing HSCT status in our sample. What was much more difficult was identifying the graduation date for former CGP households in treatment CWACs, because the records were old and had been archived (or were simply no longer available). We also directly asked households in our questionnaire when they had received their last CGP payment (in treatment CWACs), and whether they were current beneficiaries of the HSCT. These data are incomplete because many households in treatment CWACs in Shangombo and Kalabo had not yet been notified of their status since retargeting had just been completed or was ongoing. Similarly, households that had graduated from the CGP for some time may not remember the exact date of their last payment.

Our strategy is to use the CGP graduation date obtained from programme records in the first instance. If that was missing (N=233) or seemed wrong (for example there were seven cases with graduation dates just a few months after the programme began) we tried two replacement approaches: 1) We used the expected or simulated graduation date based on the birth date of the focal child; 2) we used the self-report of graduation from the household survey (starting in 2017 and going back in time to 2014 and then 2013) and where this was also missing, then the simulated graduation date of the focal child. We compared the resulting distribution of graduation dates with self-reports of graduation dates from the past household surveys, and found that the first approach best represented the pattern of graduation as reported in prior survey rounds. In the 36-month survey 11 percent of treated households self-reported that they had graduated, while in the 48-month survey 30 percent said they were no longer in the CGP.

Figure 2 shows the breakdown of HSCT-eligible and ineligible in our original CGP study arms based on the balanced panel of households that appear in all study waves. Appendix A provides the flow chart and sample sizes for all households that appear in any wave. The HSCT eligibility rate in the sample is 21 percent, so given the eligibility criteria of the new harmonised programme, most CGP households did not qualify. Those in the control arm that did not qualify received a lump-sum payment of ZMW 500 at the same time as the first HSCT payment date in their district as compensation for participating in the study. As can be seen in Figure 2, this lump-sum payment occurred over one year before our survey was fielded.

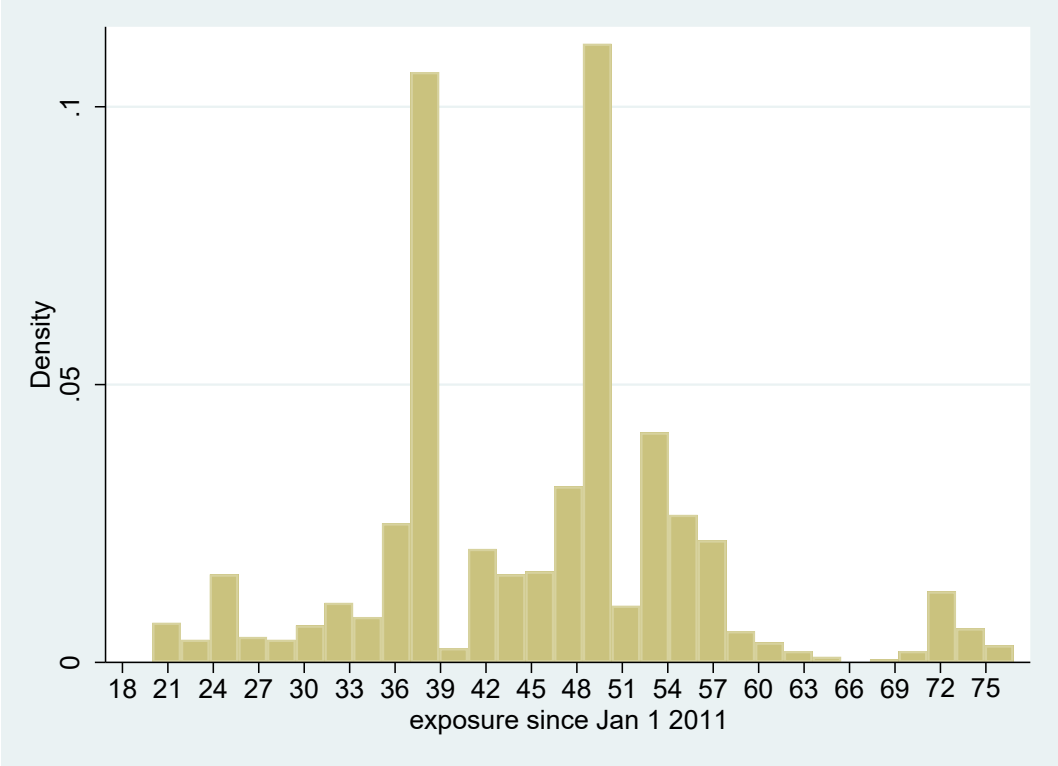
**Figure 2: Flowchart of eligible and ineligible households in evaluation sample (balanced panel)**




Note: Ns refers to the balanced panel of 2,109 households used in this report and are based on three waves of data: baseline, 36-month wave and 84-month wave. Information on HSCT beneficiary status is missing for 5 households.

Figure 3 shows the exposure to the CGP among treated households (N=1,033) in months from the first payment date in January 2011. There is a mode at 36-40 months which corresponds to children at the upper age range of 2.5-3 years old reaching age six and thus graduating; there is a second, larger mode at 48-50 months, or four years after the first payment. Our understanding from talking to the district social welfare officers is that graduation was not implemented continuously or automatically, which would explain the modes in Figure 3. *The variation in graduation dates (or exposure) is clearly important for our empirical analysis, as exposure would affect time since intervention and thus fade-out.*

**Figure 3: Exposure to the CGP in Months**




### 3.4 Empirical Approach and Statistical Specification

As indicated in the pre-analysis plan, we replicate the statistical specification employed in Handa et al (2018) with the addition of the 2017 (84-month post CGP baseline) survey data. Our most common specification is a difference-in-differences (DD) using the three survey waves of baseline (2010), 36-months (2013) and the most recent 2017 survey (84-months):

(1) $$Y_{ijt} = \beta_0 + \beta_1 T_{ij} + \beta_2 2013_t + \beta_3 2017_t + \beta_4 T_{ij} * 2013_t$$
$$+\beta_5 T_{ij} * 2017_t + \beta_6 X_{ij} + \mu_j + \varepsilon_{ijt}$$

In equation (1), $T_{ij}$ is a dummy indicating whether the household belongs to the original treatment group in the CGP evaluation, 2013 and 2017 are indicators of the survey round. $Y_{ijt}$ is the outcome of interest for the ith household in CWAC j in time period t (2010, 2013, or 2017), and X is a vector of covariates that are all measured at baseline.

The timing of the CGP graduation and the enrolment in the HSCT varies by household, but at the time of the 2017 survey, we can identify four groups:

Group 1: T Eligible (CGP and HSCT);
Group 2: T Ineligible (CGP and No-HSCT);
Group 3: C Eligible (No-CGP and HSCT);
Group 4: C Ineligible (No-CGP and No-HSCT).

T and C refers to the original assignment status in the CGP evaluation and eligible and ineligible refers to the HSCT programme eligibility status.

Our main hypothesis, fade-out, is tested by using the HSCT ineligible groups and comparing outcomes between original treatment and control households (groups 2 and 4). The coefficient $\beta_4$ provides the impact of the CGP in 2013—this is a 'replication' of the results published in Handa et al (2018). The key question is whether this impact persists or is eliminated, which is a test of $\beta_4 = \beta_5$. The standard significance test on $\beta_5$ is also of interest as it provides an indication of whether there is growth in the original control group or whether the original treatment group reverted back to pre-programme levels.

Our secondary hypothesis of catch-up replicates equation (1) using groups 1 and 3, those who are eligible for the HSCT. The interpretation of the coefficients is analogous to what is described above. All impact estimates are intent-to-treat effects (ITT).

We also perform three variations to the main analysis using equation (1) to capture the differential exposure to the CGP among original T households. First, we estimate equation (1) on a sub-set of T households that were exposed for 60+ months (high exposure) and then a sub-set that were exposed for more than the median exposure time (48 months). We expect that in these sub-samples we might find less fade-out. Finally, we use a continuous exposure variable (in logs) where C households get a value of 0, and T households get their actual (log) exposure time in months. This variable replaces $T_{ij}$ in equation (1).

There are three further issues to note in the empirical specification. First, all regressions are adjusted for a sparse set of pre-treatment variables: household demographic composition and size (logged), respondent's age, education and marital status, and district dummy variables (the stratification variable), and for child outcomes, we also included the age and sex of the child. This is the exact specification in Handa et al (2018). Second, for two of our domains, 'incomes and revenues' and 'finance and debt', we do not have baseline measures for some of the individual components: loans outstanding, amount borrowed and owed, engagement in non-farm enterprise and revenue from non-farm enterprise. In Handa et al (2018) we estimated single difference impacts for these domains using just the 24-and 36-month data. This is defensible because as we show below, and in that paper, the randomization was successful and there was very strong baseline balance across the two arms. In this study, we use both the 36- and 84-month data in a DD model, with the 36-month wave serving as the baseline. This allows us to test the fade-out and catch-up hypotheses by comparing the 36-month impact with the 84-month impact. Third, in Handa et al (2018) we show that programme effects of the CGP do not differ between the 24- and 36-month waves. We thus use 2013 as our starting point to measure fade-out, as graduation began to occur in earnest after that date, and the reported multiplier calculations are based on impacts at 36-months.

## 3.5 Ethics Approvals

The UNC IRB and the University of Zambia Ethics Committee approved the study protocols.

# 4. Findings

## 4.4 Other Operational Concerns

A potential concern with cash transfer programmes is that the beneficiary's family members may ask the recipient to shoulder extra responsibility by caring for additional people, thus decreasing the potential impact of the cash transfer on the originally targeted household. However, only four percent of programme participants reported such requests. We conclude that the programme impacts are not affected by familial solicitation.

Similarly, it is possible that local shopkeepers may instate surge pricing on programme payment days. However, fewer than 10 percent of beneficiaries said shopkeepers raise prices on payment days. Analysis from previous survey rounds showed no differential trend in prices across treatment and control CWACs, suggesting that local price inflation generated by the cash transfer did not adversely erode the value of the transfer. A similar analysis with the 84-month data again indicated no differential inflation across the original treatment and control CWACs, although that comparison is less relevant as the HSCT had reached all control CWACs in the study sample.

Another potential concern involves safety. Nearly all recipients (99 percent) keep their cash at home; usage of savings schemes outside the house is not common in our sample. As Table 5 below shows, over 40 percent of participants reported concerns about the safety of keeping their cash at home.

**Table 2: Do you feel concerned about the safety of the cash when you receive it?**

| | **Frequency** | **Percent** |
|---|---|---|
| Yes | 130 | 40.63 |
| No | 190 | 59.38 |
| *Total* | 320 | 100.00 |

However, as Table 6 demonstrates, these safety concerns do not appear to be related to intra-household conflict; more than 88 percent of participants reported no household disagreements about the cash benefits (Table 6). It is possible that intra-household conflict is underreported, but it is likely that households' safety concerns stem from worries about theft.

**Table 3: In the last 12 months, how many times have there been disagreements between household members?**

| | **Frequency** | **Percent** |
|---|---|---|
| Never | 283 | 88.44 |
| Once | 16 | 5.00 |
| 2-3 times | 16 | 5.00 |
| More than 3 times | 5 | 1.56 |
| *Total* | 320 | 100.00 |

## 4.1 Programme Knowledge

In 2017 we administered a short module on programme participation, modified from the more detailed module used in previous waves when the CGP was operating. There are 320 households who reported that they were current beneficiaries of the HSCT, about 15 percent of the sample. Programme data indicates that 21 percent of the original CGP sample qualified for the HSCT, but many households in Kalabo and Shangombo would have not yet been told of their eligibility status as retargeting was just being completed in those areas at the time of our survey.

## 4.2 Perceived Conditions

Although the HSCT is explicitly unconditional, 71 percent of our sample of 320 believe there are rules that are attached to the benefits. Those who believed there were rules were asked to name up to three rules, and the most common rules are reported in Table 2. Among those who believed there are requirements to continuing eligibility, the most commonly cited rules are those listed in Figure 4 below, and consist of school enrolment for children, providing food to young children, and investing in business or livestock.

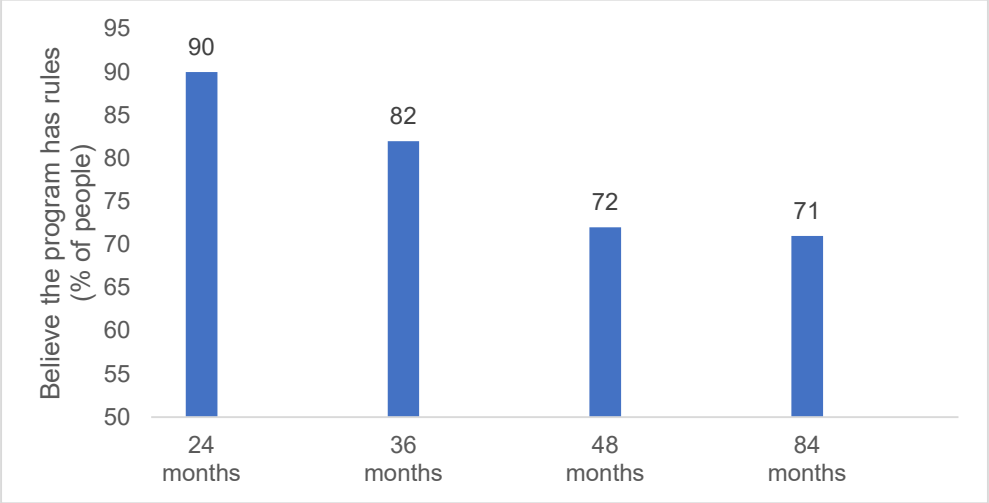
**Table 4: Types of rules or conditions that must be followed**

| Rule | Frequency | Percent of respondents who reported this rule |
|---|---|---|
| Enrolment/attendance at school | 128 | 40.00 |
| Provide food for young children | 108 | 33.75 |
| Invest in business | 75 | 23.44 |
| Invest in livestock | 73 | 22.81 |

To assess the progress of programme understanding over time, we compared the percentage of beneficiaries who believe there are continuing eligibility rules in prior survey waves. In wave 2 (the 24-month follow-up of the CGP), almost 90 percent of recipients thought they had to maintain certain eligibility requirements to continue receiving benefits. By wave 4 (the 48-month follow-up), approximately 72 percent of participants thought the programme was conditional. The 84-month follow-up shows that nearly 71 percent of beneficiaries believe there are programme rules for continuing eligibility. Programme understanding around rules or conditions seems to have plateaued after its initial increase between 2012 and 2014. It is likely that the continued confusion about conditionality is linked to the transition between the CGP and HSCT programmes; beneficiaries might assume that the new programme has rules, even though the old programme did not.[4]

[4] During the main CGP impact evaluation between 2010-2014, we tested whether program impacts varied by perception of conditions and found no differential effects.

**Figure 4: Perceived Conditionality Over Time**




## 4.3 Payments

The HSCT pays recipients every two months. In our sample, 96 percent of beneficiaries had received their most recent payment on time, an improvement of five percentage points since the 30-month follow-up survey in 2013. As Table 3 shows, approximately 83 percent of recipients felt confident they would receive their next payment as regularly scheduled. This result represents high confidence in the government and its implementation of the HSCT.

**Table 3: When do you expect to receive the next payment?Table 5: When do you expect to receive the next payment?**

| | Frequency | Percent |
|---|---|---|
| In the next two months | 266 | 83.13 |
| In the next six months | 24 | 7.50 |
| In the next twelve months | 9 | 2.81 |
| In more than twelve months | 4 | 1.25 |
| Never | 17 | 5.31 |
| *Total* | 320 | 100.00 |

Furthermore, as Table 4 shows, over 91 percent of participants reported confidence that they would continue to receive benefits five years or longer, an increase of seven percentage points from the 24-month survey. High confidence in programme delivery demonstrates that the HSCT can allow recipients to count on the money, and to make future plans based on a change in their permanent income.

**Table 6: How long in the future do you expect to continue receiving money?**

| | Frequency | Percent |
|---|---|---|
| 6 months | 2 | 0.63 |
| 1 year | 4 | 1.25 |
| 2 years | 22 | 6.88 |
| 5 years | 106 | 33.13 |
| Longer/for the rest of life | 186 | 58.13 |
| *Total* | 320 | 100.00 |

## 4.5 Attrition and Balance

The study flow chart is provided in Appendix A/B. The original sample at baseline was 2,519 households of which 2,098[5] were located and interviewed in 2017, an attrition rate of 17 percent. Field notes from the enumerators indicated the main reasons for attrition or non-contact were migration, record-keeping due to the formation of new CWACs and districts, and household dissolution. There is no differential attrition across original T and C samples. Attrition is highest in Shangombo at 21 percent, followed by 17 and 13 percent in Kalabo and Kaputa respectively.

We begin our attrition analysis by investigating *overall attrition*—whether our sample in 2017 is representative of the original baseline sample. This essentially entails comparing mean differences between the attriters and the panel on baseline characteristics. Table 7 and 8 show these means tests for the background variables that we use in our regression models as controls plus a handful of other characteristics. There are two significant mean differences across these two tables (number of children age 0-5 in the household, and the age of the recipient or potential recipient). The effect size shows that neither of these differences are more than 0.2 standard deviations in magnitude.

Additional attrition tables covering over 50 variables (all measured at baseline, as is necessary for this test) are presented in Appendix E for the reader to assess. Statistically significant differences are shown in bold—while there are a handful of these, no effect size is near 0.2 SD, and in some cases statistical significance is driven by very low overall means (e.g. the mean for owning ducks is just 1 percent). We conclude that overall, based on a comparison of 62 variables covering demographics, assets, housing and expenditures, overall attrition is not an issue in our sample and that our longitudinal sample is representative of the original sample.

5 Our final analysis sample is 2109, as we had missing data on 11 households in the balanced panel.

**Table 7: Household-level Characteristics (Attriters versus Panel Households)**

| Variables | Attriters Mean | N1 | Panel Mean | N2 | Mean Diff | Diff SE | p-value | Effect Size |
|---|---|---|---|---|---|---|---|---|
| Household size | 5.482 | 421 | 5.737 | 2,098 | 0.255 | 0.133 | 0.058 | 0.121 |
| Number of people ages 0 - 5 | 1.800 | 421 | 1.923 | 2,098 | 0.123 | 0.042 | 0.004 | 0.159 |
| Number of people ages 6 - 12 | 1.178 | 421 | 1.281 | 2,098 | 0.103 | 0.068 | 0.135 | 0.090 |
| Number of people ages 13 - 18 | 0.508 | 421 | 0.572 | 2,098 | 0.064 | 0.043 | 0.144 | 0.076 |
| Number of people ages 19 - 35 | 1.401 | 421 | 1.312 | 2,098 | -0.089 | 0.049 | 0.070 | -0.107 |
| Number of people ages 36 - 55 | 0.487 | 421 | 0.551 | 2,098 | 0.065 | 0.039 | 0.102 | 0.089 |
| Number of people ages 56 - 69 | 0.086 | 421 | 0.068 | 2,098 | -0.018 | 0.016 | 0.276 | -0.064 |
| Number of people ages 70 or older | 0.021 | 421 | 0.030 | 2,098 | 0.008 | 0.008 | 0.334 | 0.047 |

Note: Standard errors obtained by clustering at the CWAC-level

**Table 8: Main Respondent/Original CGP Recipient Characteristics (Attriters versus Panel Households)**

| Variables | Attriters Mean | N1 | Panel Mean | N2 | Mean Diff | Diff SE | p-value | Effect Size |
|---|---|---|---|---|---|---|---|---|
| Age of recipient | 28.788 | 419 | 30.060 | 2,093 | 1.272 | 0.522 | 0.017 | 0.133 |
| Gender of recipient male | 0.007 | 419 | 0.010 | 2,093 | 0.003 | 0.005 | 0.570 | 0.030 |
| Recipient ever attended school? | 0.735 | 419 | 0.724 | 2,092 | -0.011 | 0.024 | 0.651 | -0.024 |
| Recipient is disabled (0/1) | 0.014 | 419 | 0.010 | 2,093 | -0.005 | 0.006 | 0.417 | -0.047 |
| Never married | 0.083 | 421 | 0.112 | 2,098 | 0.028 | 0.021 | 0.178 | 0.092 |
| Divorced | 0.069 | 421 | 0.073 | 2,098 | 0.004 | 0.014 | 0.773 | 0.016 |
| Widow | 0.057 | 421 | 0.066 | 2,098 | 0.009 | 0.012 | 0.468 | 0.036 |
| Married | 0.748 | 421 | 0.716 | 2,098 | -0.032 | 0.029 | 0.265 | -0.072 |

Note: Standard errors obtained by clustering at CWAC-level

We now turn to *differential attrition* (or balance) across treatment and control arms, a potentially more serious concern as it affects internal validity. Here we must remember that while assignment to CGP was done randomly in 2010, we are now comparing sub-sets of the original treatment and control arms, those sub-sets are households that are eligible (or not) for the HSCT. And since there is a selection rule that includes demographic characteristics, we can expect that households selected for the HSCT will be different from those who were not. However, we can also expect that all households selected (or not) for the HSCT would be similar, irrespective of their original treatment status in the CGP, provided that the eligibility rules were applied consistently and did not change significantly over time.

We began the analysis of differential attrition (and thus balance) by comparing the original T and C samples for the CGP evaluation. We found just three statistical difference in means out of the 62 comparisons (goats, number of members age 13-18 and whether main respondent was divorced). None of these differences approached 0.2 SDs.[6]

The more important comparison for our analysis is balance within our HSCT-eligible and ineligible groups, as these are the two groups we compare. For these groups we again compared mean differences in the same 62 variables (all measured at baseline). Results are provided in Appendix F for the HSCT-ineligible groups and Appendix G for the eligible groups. For our main sample of interest, the HSCT-ineligible group, none of the 62 mean comparisons yield a statistically significant difference, allowing us to conclude that selective or differential attrition is not a concern, and that balance exists between the two groups.

Among the HSCT-eligible group there are four statistically significant mean differences and these magnitudes are greater than 0.2 SD. In one case (goats) the mean is very low, just 3.5 percent in this sample. Recall that just 21 percent of the originals sample qualified for the HSCT so the overall sample is small, just 487 households, which could be contributing to some large mean differences. One of these variables (number of residents age 13-18) is controlled in the analysis. Two variables, ownership of goats and axes are outcomes that feed into the asset index and the fourth (ownership of mobile phone) is not used in the analysis.

To illustrate the selection process in the HSCT we also compared mean differences between HSCT-eligible and ineligible households within each original arm. As we would expect, these households are quite different, with over 20 statistical differences in means across the 62 variables considered.

## 4.6 Descriptive Analysis

Given our study design, and in particular the strong balance across the T and C ineligibles and the fact that the HSCT enrolment is supply-driven and thus not dependent on self-selection, we do not need heavy statistical machinery to produce unbiased estimates of the effects of interest. To this end Figure 5 shows the trends in some of the key indicators for the T and C ineligible to see if fade-out or catch-up is visible. In each

[6] These results are available upon request: shanda@email.unc.edu

graph, the left bars are the original C households and the right bars are the original T households; the name of the indicator is shown above the graph. Given the attrition and balance analysis we expect the bars representing 2010 (baseline) to be the same height across the two samples, followed by a sharp increase in the treatment sample in 2013 relative to the control sample. Of interest to this study is the trend between 2013 and 2017 in the two groups, which would suggest fade-out, catch-up, or sustained impacts.

The top left graph in Figure 5 shows trends in consumption per capita. There is clear convergence in consumption by 2017—consumption in the original T group drops off in 2017 while it continues to grow in the original C group. This basic pattern is repeated for food consumption and the summary indicator for food security (third graph in row 1). This basic pattern holds for the productive activity index, the single indicator of savings which is part of the 'finance and debt' domain, and for the two subjective well-being indicators which form part of the relative poverty/subjective well-being domain. Indeed, it is telling that there is a large drop in the proportion of T households who say their life is better than a year ago while there is a large increase in the C group. This pattern is consistent with the well-known idea that individuals need to be compensated much more for losing an entitlement they currently have than they are willing to pay to obtain that same entitlement if they do not currently have it.

There are some exceptions to the fade-out or convergence pattern. There are increases in both groups in the livestock index, the value of harvest and expenditure on agricultural inputs (these last two form part of the income/revenue domain). However, in all cases the increase in the C group exceeds the corresponding increase in the T group, suggesting again a story of convergence, albeit a happier version of convergence driven by catch-up rather than fade-out.

It is important to remember that the C ineligible group received a ZMW 500 lump-sum payment at the time of re-targeting, equivalent to 6-months' worth of transfers. This lump-sum would also explain the apparent increase in consumption and other indicators since 2013.

**Figure 5: Means by study wave and original treatment status among HSCT ineligibles sample (fade-out)**

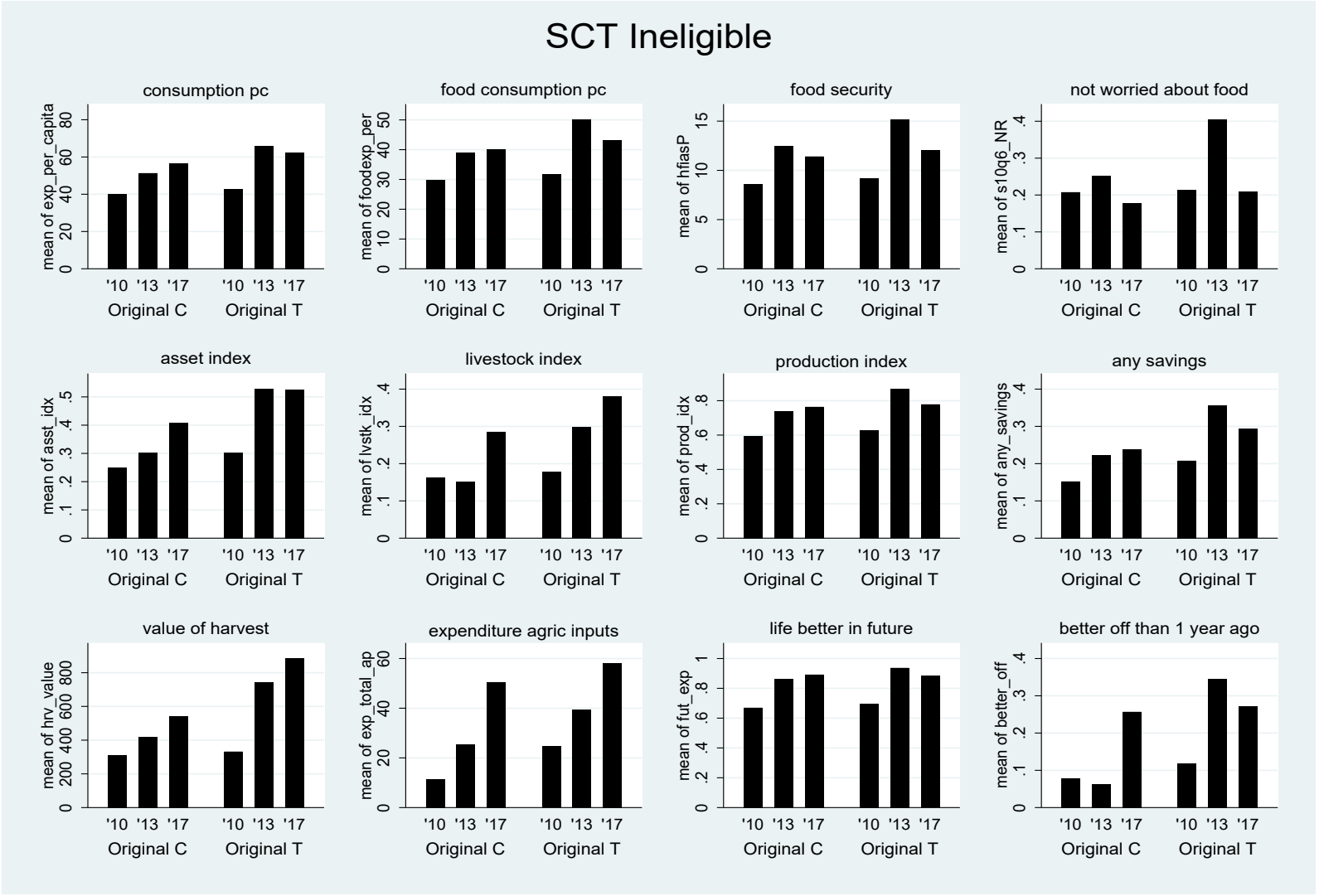


Figure 6 repeats the graphs for the HSCT eligible sample, where the pattern is not as clear cut. For example, the consumption indicators (total and food) increase in both T and C between 2013 and 2017 at an equal rate, but food security and the key indicator of worrying about food drop off at a much greater rate in the T group. There are no clear patterns in the productivity and asset indicators either. The asset index increases significantly among the C group but remains flat in the T group, while the livestock index increases in both groups and by much more in the C group. Both groups see drops in the value of harvest, but the decline is larger in the T group. Recall that the sample size for this sample is small, making it potentially harder to gain any meaningful signal from the data.

**Figure 6: Means by study wave and original treatment status among HSCT eligible sample (catch-up)**

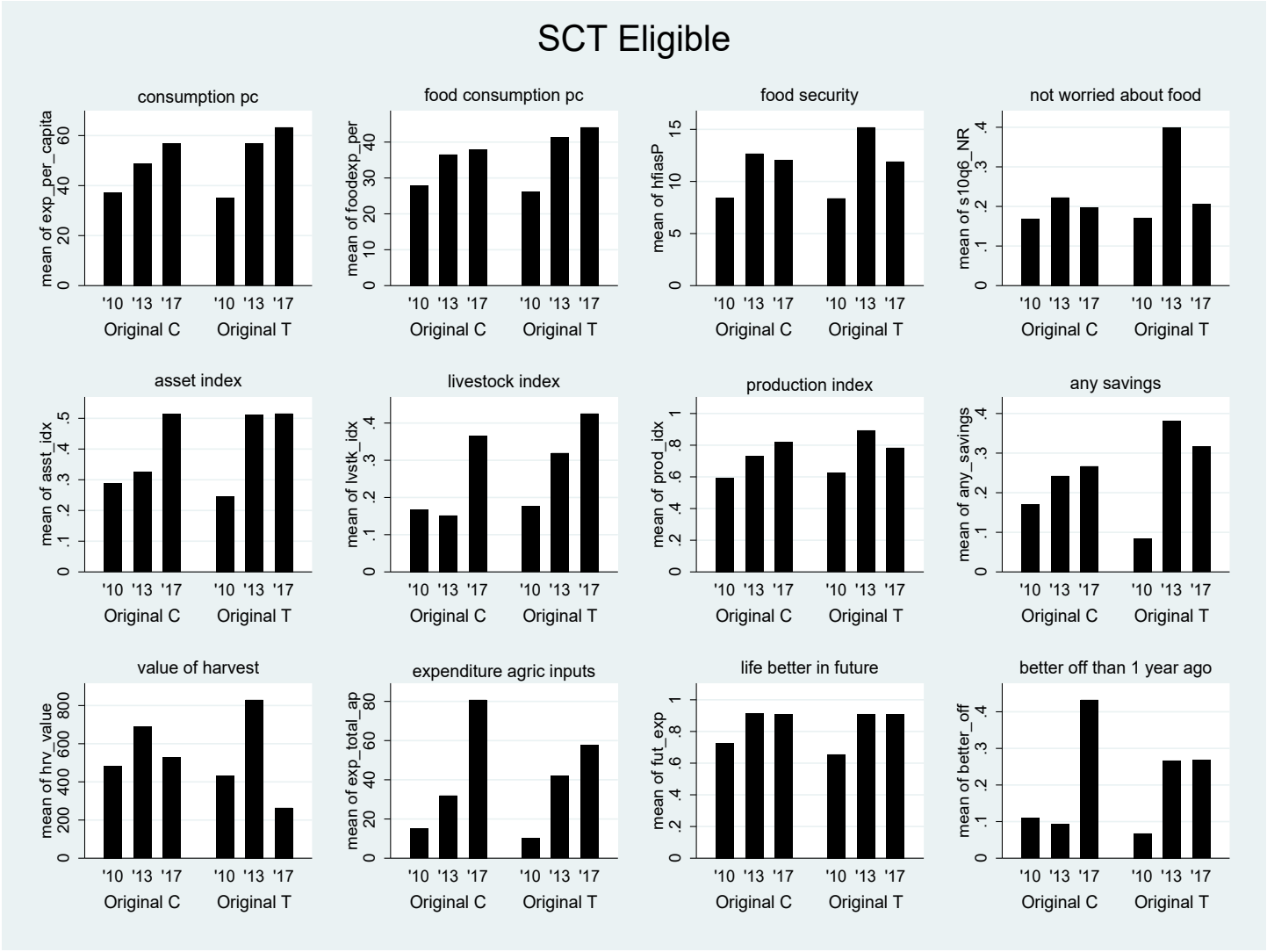
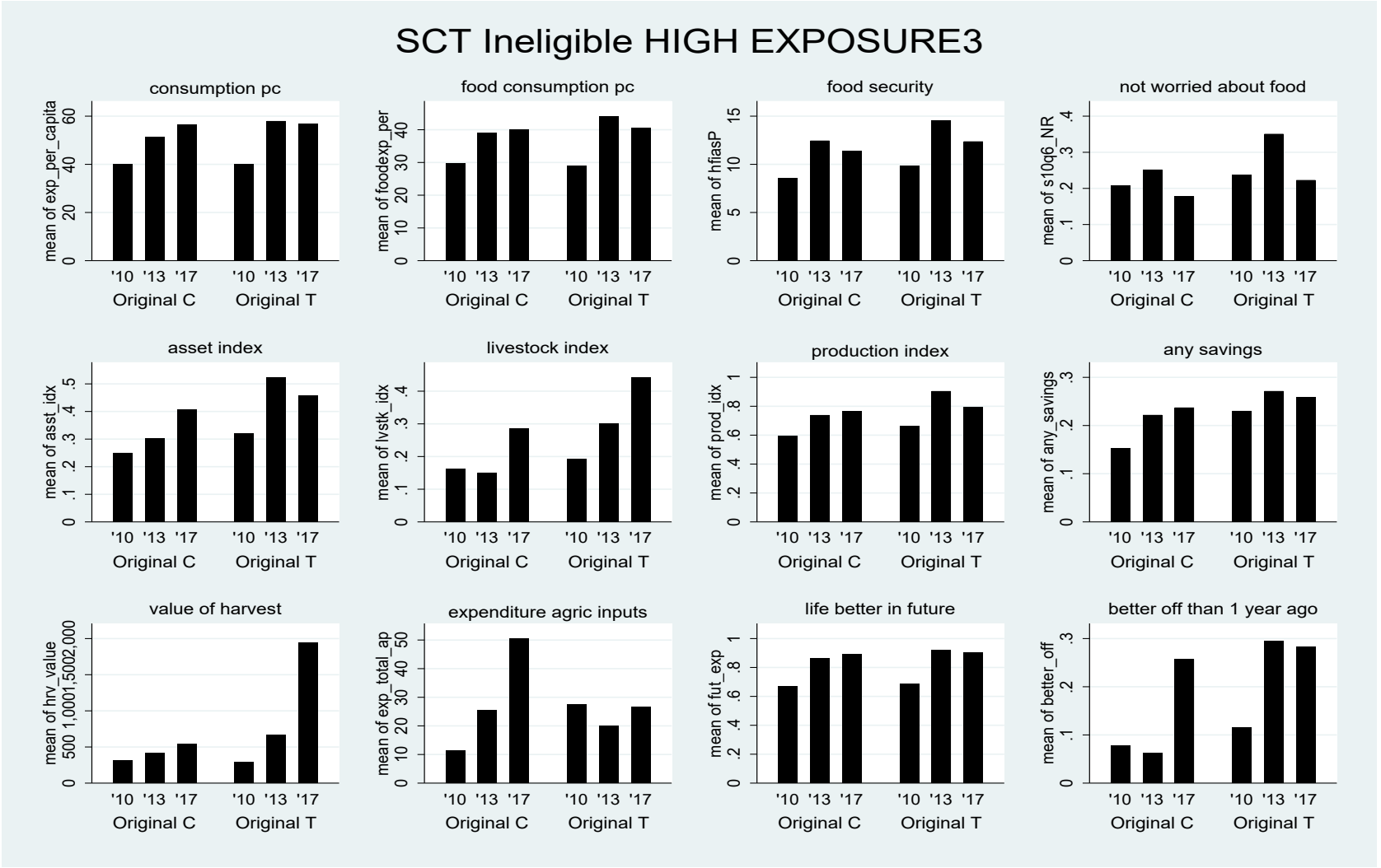


Figure 7 restricts the sample of T ineligibles to those with high exposure to the CGP (defined here as the top tercile of exposure). We might expect fade-out to be less of an issue here as T households received cash transfers for longer in the CGP. Indeed, the drop-off in consumption and food consumption between 2013 and 2017 are smaller among T households, but food security and worrying about food continue to worsen significantly, and much more relative to the original C group. The productive and asset indicators are somewhat inconsistent. Both the asset index and the value of harvest increases in the T group in 2017 and by much more than the C group; on the other hand, there are drops in the production index and asset index while there are increases in those same indexes in the C group. In general, then it might be harder to make an unambiguous argument for fade-out in this group.

**Figure 7: Means by study wave and original treatment status among HSCT ineligible sample and high exposure to CGP**



## 4.7 Impact Analysis

In this Section, we discuss whether the impacts of the Child Grant Program were sustained seven years after the cash transfers ended. First, we present 'fade-out' results which refer to the post-intervention impacts for households deemed ineligible for the new HSCT. Second, we show findings related to that minority of households who were eligible to the new HSCT programme. In all of the tables reported hereafter the main focus is on the p-value of the test between the 36- and 84-month impacts, as already highlighted in the identification strategy section.

Tables of means by wave and treatment status for all outcomes of interest (in actual units) are reported in Appendix H. Tables H.1-H.3 refer to HSCT ineligible households whereas tables H.4-H.5 refer to HSCT eligible households only. All tables include simple DDs (programme impacts) at midline (36 months) and endline (84 months).

### *4.7.1 Fade-out: Long-term impacts of the CGP among Harmonized SCT ineligibles*

Here, our main interest is to understand whether the strong protective and productive impacts of the CGP - recorded while the programme was still on-going - were sustained or faded out over time as beneficiary households stopped receiving their transfers.

Table 9 report impacts of the CGP for our largest sample: 1,615 households that were not deemed eligible for the Harmonized SCT (or roughly 77 percent of our initial evaluation sample). Three years into the programme, we confirm the broad impacts of the programme already reported in Handa et al. (2018) on almost every single domain (see DD[36-Month] coefficients for columns 1-7 and 10-11 and Treatment coefficients for specifications 8 and 9). Coefficients are large and strongly significant. Total per capita consumption increased by 0.39 standard deviation units of the control (col.1) which is equivalent to a 20 percent impact off the control group mean at baseline; the impact on overall consumption is mainly driven by an increase in food consumption. In line with these findings, there is a positive impact on the food security index of 0.47 SD (col. 2) or 24 percent of baseline control mean.

The overall asset index increased by 0.5 SD (col. 3). The impact on each of the sub-components of the overall asset index is highly statistically significant indicating that the programme did not only have a protective impact (+0.49 SD for the domestic assets index, col. 4) but also enabled households to invest in livestock and productive assets (0.49 SD and 0.29 SD impact for the livestock and productive indexes respectively, cols. 5 and 6). Raw means (in actual units, not z-scores) by wave and treatment status for the HSCT ineligibles - including simple DDs – are presented in Appendix H (Tables H.1-H.3): here we also report means for any asset included in each index[7].

Beneficiary households were also better off in terms of their overall incomes and revenues, as well as their financial situation. We do not have baseline data for all indicators in these two domain indexes and so, for these specifications (cols. 8 and 9), our coefficient of interest is the treatment dummy ('Treatment') which captures the difference between treatment and control at 36-month. The magnitude of the impact is

7 This table for instance indicates that larger impacts at midline were recorded for ownership of bed and mattresses among domestic assets and for chickens and cattle among livestock.

0.39 and 0.25SD for the ‘incomes and revenues’, and the ‘finance and debt’ indexes respectively. Overall, subjective measures of household wellbeing captured by the relative poverty index also improved by almost 0.8 SD (col. 7). Finally, although there does not seem to be an overall impact on children’s schooling, the programme helped households covering for their children’s material needs; the strongly significant impact is over 0.5 SD equivalent to roughly 27 percentage points[8] or a threefold percent increase over baseline control mean; these impacts are driven by a positive impact on ownership of shoes and blankets whereas there is no significant impact on owning two sets of clothing.

By endline (84-month wave), and after the programme ended, all the effects on protective and productive indexes are no longer significant (not shown for ‘incomes and revenues’ and the ‘finance and debt’ indexes). The magnitude of impacts is always lower than those reported at 36 months. Indeed, the differences between the 36- and 84-month impacts are highly statistically significant ($p<0.01$ as reported at the bottom of the table, $p<0.05$ for ‘finance and debt’) for all domains but schooling; this was the only domain not found to be significant at 36 months either[9].

Evidence suggests the impacts of the programme completely faded away in every single domain indicating full convergence of the control group by 84 months.

8 Results are broadly consistent with those reported in Handa et al. (2018). Any discrepancy is mainly related to the use of a slightly different sample (focus here is on the balanced panel using baseline-36m-84m waves and on ineligibles only).

9 As noted in Handa et al. (2018), and Handa et al. (2016), there is a positive impact on schooling when the analysis focuses on children 11-14 years old. This is mainly a consequence of the CGP eligibility criteria which targets young households with few adolescents aged 14-17 years old. We therefore replicated the impact of the CGP on whether or not the child was attending school (binary, raw indicator) and find a 5 percentage points impact at 36 months ($p=0.08$) and a 6 percentage point impacts at 84 months ($p=0.1$) approaching marginal significance.

**Table 9: Effects of being a beneficiary of CGP on domain indices among HSCT ineligibles**

| | (1) Total consumption per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Productive asset index | (7) Relative poverty index | (8) Incomes & Revenues index* | (9) Finance & Debt index* | (10) Schooling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | 0.0324 | 0.0684 | 0.0468 | 0.0428 | 0.0208 | 0.0402 | 0.130 | 0.391*** | 0.253*** | -0.0225 | -0.0147 |
| | (0.0641) | (0.104) | (0.0727) | (0.0754) | (0.0735) | (0.0709) | (0.0857) | (0.0896) | (0.0950) | (0.0763) | (0.0695) |
| DD[36-Month] Treatment*36-Month | 0.387*** | 0.473*** | 0.544*** | 0.488*** | 0.486*** | 0.287*** | 0.790*** | | | 0.0948 | 0.529*** |
| | (0.0850) | (0.132) | (0.104) | (0.101) | (0.0903) | (0.0981) | (0.107) | | | (0.0796) | (0.103) |
| DD[84-Month] Treatment*84-Month | 0.0391 | 0.0433 | 0.0824 | 0.113 | 0.137 | -0.0537 | -0.141 | -0.266* | -0.179 | 0.0584 | -0.0528 |
| | (0.112) | (0.130) | (0.101) | (0.114) | (0.0885) | (0.0970) | (0.117) | (0.138) | (0.122) | (0.0841) | (0.0893) |
| Observations | 4,842 | 4,786 | 4,844 | 4,840 | 4,804 | 4,830 | 4,809 | 3,230 | 3,230 | 4,866 | 11,990 |
| R-squared | 0.190 | 0.071 | 0.192 | 0.184 | 0.099 | 0.124 | 0.136 | 0.085 | 0.019 | 0.073 | 0.118 |
| p-value (36m=84m) | 0.001 | 0.002 | 0.00 | 0.00 | 0.00 | 0.001 | 0.00 | 0.003 | 0.042 | 0.604 | 0.00 |

Note: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### *4.7.2 Do CGP impacts vary by intervention exposure among Harmonized SCT ineligibles?*

We now turn to see whether the impact of the CGP varies by the length of exposure to the intervention. We first report estimates from a dose-response model where the treatment level or the dose is defined as the months the household has received cash since the beginning of the first transfer payment in January 2011. We then report intent-to-treat effects for households that were exposed the longest to the programme using two thresholds: more than 48 or 59+ months. One possible concern with these estimates is if households could somehow influence their exposure time in the CGP. If more productive and industrious households somehow were able to increase their exposure time, they might also use the cash transfer more productively and realize higher gains in consumption, which might lead to persistent effects of the CGP among those with higher exposure. As explained earlier, graduation from the CGP occurred when the focal child turned 6 years of age. Birth dates were recorded at the time of enrolment into the programme, and graduation (based on these birth dates) was implemented inconsistently across districts, and seemed to have occurred about once or twice a year when the district social welfare office reviewed birth dates. We checked to see if there was any correlation between household characteristics (other than age of focal child) and length of exposure and found no relationship between exposure and baseline consumption or assets such as livestock.

**Dose-response approach**

Table 10 report the impacts of exposure to the CGP on each domain index among HSCT ineligibles only. Exposure among HSCT ineligibles and original CGP beneficiaries ranges between 20 and 77 months (whereas it is 0 for control HSCT ineligible households). We report again, in line with the previous Section, highly significant impacts after 3 years and while the programme is still on-going. Impacts tend to increase with exposure to the programme. On average, a 10 percent increase in exposure results at 36 months in: a 0.01 SD increase in total consumption per capita, a 0.012 SD improvement in food security, a 0.014 SD rise in overall asset index (the increase on each individual sub-components varies between 0.008 and 0.02), a 0.13 SD increase in children's material needs, a 0.01 SD increase in the incomes and revenues index (see exposure coefficient in col.8) and a 0.006 SD improvement in households' financial situation (see exposure coefficient in col.9). The interaction term (exposure*follow-up) is again not significant for the schooling index in line with previous results.

Moving now on to the findings at 84 months, there is no evidence that an increase in exposure raises the likelihood of finding an impact on any of the domains observed. None of the interaction terms is ever even marginally significant (not shown for col. 8 and 9) and the size of the coefficients is also lower compared to the 36-month estimates. Indeed, 36- and 84-month treatment level impact estimates are statistically significantly different from each other in all domains (see p-value reported at the bottom of the table) apart from schooling (p-value>0.1). These results are consistent with those shown in the previous Section. We have also repeated these estimates using only households where the graduation date is reported (removing households where we simulated the graduation date) and found the results to consistent with those in Table 10.

**Table 10: Effects of exposure to CGP on domain indices among HSCT ineligibles (Dose-response)**

| | (1) Total consumption per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Productive asset index | (7) Relative poverty index | (8) Incomes & Revenues index* | (9) Finance & Debt index* | (10) Schooling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Exposure (logged) | 0.00673 | 0.0181 | 0.0123 | 0.0127 | 0.00557 | 0.00903 | 0.0322 | 0.100*** | 0.0652** | -0.00513 | -0.00235 |
| | (0.0165) | (0.0273) | (0.0192) | (0.0197) | (0.0193) | (0.0187) | (0.0223) | (0.0235) | (0.0248) | (0.0196) | (0.0180) |
| DD[36-Month] Exposure*36-Month | 0.1000*** | 0.122*** | 0.144*** | 0.127*** | 0.130*** | 0.0778*** | 0.205*** | | | 0.0229 | 0.134*** |
| | (0.0222) | (0.0349) | (0.0272) | (0.0263) | (0.0235) | (0.0255) | (0.0277) | | | (0.0209) | (0.0270) |
| DD[84-Month] Exposure*84-Month | 0.0132 | 0.0104 | 0.0210 | 0.0269 | 0.0360 | -0.0129 | -0.0339 | -0.0693* | -0.0447 | 0.0148 | -0.0167 |
| | (0.0292) | (0.0338) | (0.0264) | (0.0303) | (0.0233) | (0.0253) | (0.0302) | (0.0357) | (0.0318) | (0.0219) | (0.0232) |
| Observations | 4,842 | 4,786 | 4,844 | 4,840 | 4,804 | 4,830 | 4,809 | 3,230 | 3,230 | 4,866 | 11,990 |
| R-squared | 0.190 | 0.071 | 0.192 | 0.184 | 0.099 | 0.125 | 0.135 | 0.084 | 0.019 | 0.073 | 0.118 |
| p-value (36m=84m) | 0.00200 | 0.00200 | 0 | 0 | 0 | 0.00100 | 0 | 0.00300 | 0.0480 | 0.659 | 0 |

Note: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### *4.7.3 CGP impacts among most exposed households*

To try and understand whether any of the impacts might have survived after cash transfers ended for some sub-groups of the population, we focus on households that have been exposed the longest to the programme. Tables 11 and 12 report results for households who received the programme for at least four and five years respectively. These cut-offs, as already highlighted, were selected on the basis of the exposure distribution histogram among ineligibles (Fig. 3).

In line with earlier findings, Table 11 indicates that impacts at 36 months are highly significant with the only exception the impact on schooling for children aged 11-17 years old which was never found significant even in previous estimations. As we would expect, the magnitude of impacts for this group of households - who were exposed the longest to the programme - is systematically higher than those reported on the full sample of ineligibles in Table 9. No significant impact is found at 84 months; as a consequence, differences between impacts at 36 and 84 month are in most cases highly statistically significant (see p-values at the bottom of the table) with the exception of the 'finance and debt' index ($p<0.1$) and the schooling domain ($p>0.1$).

Table 12 focuses on those households that were exposed for over 5 years (>59 months). These households represent roughly 6 percent of the initial treatment ineligible households (48 out of 782 households) so the sample is small and somewhat selected, which should be kept in mind. Results indicate again strong impacts at 36 months, although the three-year impact on food security scale, incomes and revenues as well as finance and debt are not statistically significant anymore. Note that the 36-month point estimates for food security and income and revenue are actually larger than for the full sample in Table 9, so the lack of significance is driven by low power. Impacts at 84 months are not statistically significant with the exception of the impact on total consumption per capita which is marginally significant: this would most likely not survive and adjustment for multiple hypothesis testing. The magnitude of the 84-month impact is larger than the 36-month one (0.44 SD vs 0.38 SD), however impact estimates do not differ significantly from each other. The same holds for most of the other domains: as shown by the p-value at the bottom of the table, the only impact estimates that differ significantly between the 36-month and the 84-month waves are those for the relative poverty index ($p<0.05$) and the domestic asset index; however the p-value is close to being marginally significant for the overall asset index and the material needs index.

As a further robustness checks, we have also run the same set of four estimations reported up till this point but focusing only on the two original districts in Western province (Kalabo and Shangombo). As explained earlier, the CGP was phased out in early 2017 in these districts whereas in Kaputa the last CGP payment took place in early 2015. This means that households who had not graduated out of the programme received the programme for far longer in Western Province than in Northern Province. Results are reported in Appendix I, Tables I.1-I.4. The first two tables - focusing on the impact of either participating in the programme or exposure to the programme - confirm large and highly significant impacts at 36 months. Here the only exceptions are that the impact on productive assets and 'finance and debt' are not statistically significant anymore. None of the impact estimates is significant at 84 months in either table indicating complete fade out of the programme once the cash transfers end also in Western Province where in theory the CGP ended later on; indeed, impacts at 36 and 84

months are statistically significantly different whenever the 36 month impact is significant with the exception of the livestock and ‘incomes and revenues’ indexes ($p<0.1$). Note of course that although the last CGP transfer in Western Province was made later on it does not mean that all treatment households were still receiving the programme; indeed, already at 48 months almost 30 percent of households were not receiving the programme anymore and over time graduation is expected to have been implemented in a stricter way. Tables I.1-I.2 report ITTs for most exposed households (more than 4 and 5 years respectively); once again, apart from schooling, impacts at 36 months are strong and significant even though with few exceptions (food security, ‘incomes and revenues’ and ‘finance and debt’ lose significance or are just marginally significant). None of the impacts at 84 months is significant using either cut-offs with the only exception of a strong and significant sustained impact of 0.65SD for households in Western province who were covered for five years or longer. For these households, the programme still has a sustained impact on this protective measure. In most cases the difference between the 36-month and the 84-month impacts is not statistically significant.

**Table 11: Effects of being a beneficiary of CGP on domain indices among highly exposed (>48 months) HSCT ineligibles**

| | (1) Total consump-tion per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Productive asset index | (7) Relative poverty index | (8) Incomes & Revenues index* | (9) Finance & Debt index* | (10) Schooling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | -0.0145 | 0.0774 | 0.0120 | 0.0572 | -0.00778 | -0.0223 | 0.0610 | 0.439*** | 0.265** | -0.0657 | -0.0211 |
| | (0.0705) | (0.121) | (0.0939) | (0.0955) | (0.0841) | (0.0881) | (0.0938) | (0.108) | (0.108) | (0.0849) | (0.0822) |
| DD [36-Month] Treatment* 36-Month | 0.430*** | 0.421*** | 0.675*** | 0.597*** | 0.558*** | 0.408*** | 0.816*** | | | 0.0930 | 0.546*** |
| | (0.0995) | (0.159) | (0.122) | (0.113) | (0.110) | (0.110) | (0.127) | | | (0.0991) | (0.121) |
| DD [84-Month] Treatment* 84-Month | 0.133 | 0.0239 | 0.128 | 0.155 | 0.148 | -0.00657 | -0.0324 | -0.366** | -0.174 | 0.0619 | -0.0490 |
| | (0.122) | (0.142) | (0.117) | (0.146) | (0.102) | (0.115) | (0.108) | (0.148) | (0.141) | (0.108) | (0.101) |
| Observations | 3,679 | 3,643 | 3,680 | 3,677 | 3,651 | 3,674 | 3,660 | 2,454 | 2,454 | 3,726 | 9,242 |
| R-squared | 0.164 | 0.054 | 0.189 | 0.173 | 0.097 | 0.124 | 0.105 | 0.073 | 0.018 | 0.069 | 0.108 |
| p-value (36m=84m) | 0.00800 | 0.00500 | 0 | 0 | 0 | 0.00100 | 0 | 0.00100 | 0.0730 | 0.721 | 0 |

Notes: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

**Table 12: Effects of being a beneficiary of CGP on domain indices among highly exposed (>59 months) HSCT ineligibles**

| | (1) Total consump-tion per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Produc-tive asset index | (7) Relative poverty index | (8) Incomes & Reve-nues index* | (9) Finance & Debt index* | (10) Schooling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | -0.0724 | 0.205 | -0.0166 | 0.124 | 0.0249 | -0.195 | 0.124 | 0.197 | -0.127 | 0.0959 | -0.0536 |
| | (0.147) | (0.236) | (0.184) | (0.152) | (0.180) | (0.173) | (0.188) | (0.193) | (0.195) | (0.118) | (0.0925) |
| DD [36-Month] Treatment* 36-Month | 0.380*** | 0.203 | 0.890*** | 0.827*** | 0.607*** | 0.624*** | 0.641*** | | | -0.0841 | 0.362** |
| | (0.131) | (0.270) | (0.147) | (0.200) | (0.210) | (0.124) | (0.213) | | | (0.192) | (0.165) |
| DD [84-Month] Treatment* 84-Month | 0.443* | -0.00688 | 0.309 | 0.0729 | 0.367 | 0.269 | -0.225 | 0.0561 | 0.228 | 0.0586 | -0.0378 |
| | (0.224) | (0.255) | (0.334) | (0.337) | (0.342) | (0.247) | (0.226) | (0.258) | (0.214) | (0.162) | (0.174) |
| Observations | 2,643 | 2,623 | 2,643 | 2,640 | 2,621 | 2,638 | 2,634 | 1,762 | 1,762 | 2,618 | 6,555 |
| R-squared | 0.152 | 0.043 | 0.157 | 0.145 | 0.084 | 0.114 | 0.059 | 0.052 | 0.023 | 0.080 | 0.068 |
| p-value (36m=84m) | 0.741 | 0.275 | 0.101 | 0.0400 | 0.385 | 0.254 | 0.00 | 0.735 | 0.373 | 0.563 | 0.104 |

Notes: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### *4.7.4 Catch-up: Long-term impacts of the CGP among Harmonized SCT eligibles*

In this Section, we focus on the subsample of households from the original evaluation sample who were retargeted to receive the Harmonized Social Cash Transfer (HSCT) as shown in the flowchart (Fig. 2). These households represent roughly 23 percent of our three-wave balanced panel so any analysis on further sub-samples should be carried out with caution. These households are 'labour constrained', as per the eligibility criteria described in Section 2. Initial CGP beneficiary households retargeted for the HSCT started receiving transfers in mid-2017 in Kaputa, whereas they were still waiting to receive the first transfer in the two Western provinces. However, all the estimations, as already noted in the identification strategy section, include district dummies - which are also the stratifying indicators used for randomization.

Table 13 shows strong and highly significant impacts of the programme after three years, in line with those reported among ineligibles in Table 9. Schooling remains the only domain where no significant impacts are recorded, whereas there is only a marginally significant impact on the 'incomes and revenues' index. Impacts tend to be larger than what is found for the ineligibles for food security (0.56 vs 0.47 SD), the overall asset index (0.72 vs 0.54 SD) and its sub-components, 'finance and debt' (0.44 vs 0.25 SD) and material needs (0.81 vs 0.53SD). Consumption impact estimates at 36 months between eligible and ineligibles are almost identical (0.39SD) whereas the impact on subjective wellbeing and 'finance and debt' are comparatively smaller (0.59 vs 0.79 and 0.22 vs 0.39SD respectively).

We find no evidence that these impacts are sustained at 84 months. Indeed, none of the impacts is statistically significant; also the magnitude of these impacts is smaller when compared to 36 months and in some cases even negative. The p-values at the bottom of the table indicate that the impacts at 36 and 84 month are significantly different from each other for all domains ($p<0.01$) with the exception of consumption and 'incomes and revenues'.

Means by wave and treatment status for the HSCT eligibles are presented in Appendix H (Tables H4-H.6) together with some simple difference in difference estimations[10].

[10] These impact estimates might therefore vary slightly from those reported in the main text as they are computed using raw indicators and without any adjustment (i.e. no controls).

**Table 13: Effects of being a beneficiary of CGP on domain indices among HSCT eligible (Catch-up)**

| | (1) Total consump-tion per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Produc-tive asset index | (7) Relative poverty index | (8) Incomes & Reve-nues index* | (9) Finance & Debt index* | (10) Schoo-ling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | 0.0212 | -0.0691 | -0.0655 | -0.139 | -0.0644 | 0.0271 | -0.0633 | 0.223* | 0.444*** | 0.205* | -0.244** |
| | (0.133) | (0.114) | (0.107) | (0.111) | (0.108) | (0.117) | (0.112) | (0.116) | (0.111) | (0.104) | (0.105) |
| DD [36-Month] Treatment* 36-Month | 0.388*** | 0.557*** | 0.723*** | 0.659*** | 0.582*** | 0.404*** | 0.589*** | | | -0.0390 | 0.808*** |
| | (0.124) | (0.184) | (0.138) | (0.127) | (0.141) | (0.138) | (0.181) | | | (0.117) | (0.160) |
| DD [84-Month] Treatment* 84-Month | 0.200 | -0.0124 | -0.00593 | 0.117 | 0.0986 | -0.198 | -0.136 | -0.267 | -0.420*** | -0.0816 | 0.0898 |
| | (0.263) | (0.179) | (0.162) | (0.169) | (0.153) | (0.167) | (0.166) | (0.182) | (0.153) | (0.124) | (0.136) |
| Observations | 1,467 | 1,446 | 1,467 | 1,463 | 1,454 | 1,454 | 1,459 | 978 | 978 | 1,846 | 4,190 |
| R-squared | 0.110 | 0.075 | 0.193 | 0.160 | 0.121 | 0.126 | 0.096 | 0.078 | 0.041 | 0.073 | 0.120 |
| p-value (36m=84m) | 0.383 | 0.00 | 0.00 | 0.00 | 0.004 | 0.00 | 0.00 | 0.08 | 0.001 | 0.698 | 0.00 |

Notes: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### *4.7.5 Do CGP impacts vary by intervention exposure among Harmonized SCT eligibles?*

**Dose response approach**

In Table 14, we use a treatment level variable (months of exposure) and interact it with each wave dummy to see how exposure to the programme might have impacted on our outcome variables of interest.

In line with previous results, impacts at 36 months are highly significant with the exception of schooling (which never is) and the 'incomes and revenues' index which – as noticed also in the previous table and section - is only marginally significant. Doubling the months of exposure, namely a 100 percent increase, leads to a 0.1-0.2 SD impact in all domains apart from 'incomes and revenues' (0.05SD, $p<0.1$) and schooling ($p>0.1$).

At 84 months, once again, we find no significant impact of increasing exposure to the CGP (not shown in cols. 8 and 9) and coefficients are fairly small. The p-values at the bottom of the table confirm a significant reduction in impacts over time with the sole exception of schooling (not significant at 36 months), consumption and 'incomes and revenues'.

**Table 14: Effects of exposure to CGP on domain indices among HSCT eligible (Dose-response)**

| | (1) Total consump-tion per capita | (2) Food security scale | (3) Overall asset index | (4) Domest ic asset index | (5) Livestock index | (6) Produc-tive asset index | (7) Relative poverty index | (8) Incomes & Reve-nues index* | (9) Finance & Debt index* | (10) School-ing index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Exposure (logged) | 0.00795 | -0.0180 | -0.0142 | -0.0340 | -0.0134 | 0.00787 | -0.0166 | 0.0535* | 0.116*** | 0.0576** | -0.0616** |
| | (0.0343) | (0.0301) | (0.0283) | (0.0291) | (0.0289) | (0.0308) | (0.0298) | (0.0313) | (0.0301) | (0.0276) | (0.0280) |
| DD [36-Month] Exposure* 36-Month | 0.100*** | 0.144*** | 0.187*** | 0.170*** | 0.151*** | 0.104*** | 0.154*** | | | -0.0133 | 0.208*** |
| | (0.0314) | (0.0475) | (0.0362) | (0.0332) | (0.0364) | (0.0361) | (0.0465) | | | (0.0311) | (0.0432) |
| DD [84-Month] Exposure* 84-Month | 0.0476 | -0.00546 | -0.00429 | 0.0290 | 0.0250 | -0.0554 | -0.0337 | -0.0621 | -0.110*** | -0.0261 | 0.0195 |
| | (0.0681) | (0.0466) | (0.0427) | (0.0445) | (0.0397) | (0.0442) | (0.0431) | (0.0480) | (0.0408) | (0.0334) | (0.0353) |
| Observations | 1,467 | 1,446 | 1,467 | 1,463 | 1,454 | 1,454 | 1,459 | 978 | 978 | 1,846 | 4,190 |
| R-squared | 0.110 | 0.075 | 0.193 | 0.160 | 0.122 | 0.126 | 0.096 | 0.077 | 0.041 | 0.073 | 0.119 |
| p-value (36m=84m) | 0.353 | 0 | 0 | 0 | 0.00400 | 0 | 0 | 0.120 | 0.00100 | 0.664 | 0 |

Notes: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### *4.7.6 CGP impacts among HSCT-eligible households exposed the most*

As we did for ineligibles, we now focus only on households that were exposed the most to the programme. Table 15 report impacts for households who should have received the programme for at least 4 years; this sample represents about 37 percent of the original treatment group among eligible (92/251 households, so a very small sample size).

Three years into the programme, there are positive impacts on beneficiary households on almost every domain, consistent with all previous findings for all sub-groups. The exception is schooling and 'incomes and revenues'. The significance is also slightly weaker than for the full sample of eligible households (Table 13): impacts on consumption, food security and relative poverty index are now significant at the five – rather than one - percent level, whereas the impact on the productive asset index is only marginally significant. The magnitude on 36-month impacts does not seem larger than those reported in Table 13, however the sample is small, the coefficients likely to be less precisely estimated and differences in 36-month impacts between eligible and ineligible unlikely to be statistically significant.

As in all previous results tables, there are no significant sustained impacts 7 years after the onset of the programme. This result holds also for these households who were eligible and retargeted for the new Harmonized SCT. There is evidence of a significant reduction in impacts between 36 and 84 months with the exception of domains on which there was no impact at 36 months to start with (schooling and 'incomes and revenues') but also consumption. Even though the consumption estimate at 84 months is not statistically significantly different than the 36-month impact, the magnitude is around 0.02 SD compared to 0.3SD at 36 months.

For eligibles, we do not report results for households that were beneficiary of the programme for at least five years. This is an extremely small portion of the treatment group among eligible households (13 out of 251, 5 percent).

Finally, it is also interesting to see whether the results vary for households in Kaputa where eligible households actually started receiving cash transfers again since May 2017 (in Shangombo and Kalabo the first cash transfer was expected for January 2018 and therefore only after our data collection exercise). The sample size is small; however, we still report these results in Appendix I, Tables I.5-I.7 (we do not report impacts on households exposed for more than 59 months as there are no such households among eligibles). Impacts at 36 and 84 months are still statistically significantly different except for schooling and consumption, both when studying the effect of being a beneficiary or the impact of exposure to the intervention; there is a marginally significant impact on schooling however the impact is still statistically different than the three-year one (Tables I.5-I.6). In Table I.7, focusing on households who were exposed for at least 4 years, there are some impacts at 84 months in the overall asset index (marginally significant and driven by highly significant impact on domestic assets); however, tests of the difference of impacts at 36 and 84 months are mostly significant.

**Table 15: Effects of being a beneficiary of CGP on domain indices among highly exposed (>48 months) HSCT eligible**

| | (1) Total consump-tion per capita | (2) Food security scale | (3) Overall asset index | (4) Domestic asset index | (5) Livestock index | (6) Produc-tive asset index | (7) Relative poverty index | (8) Incomes & Reve-nues index* | (9) Finance & Debt index* | (10) Schoo-ling index (11-17 years) | (11) Material needs index (5-17 years) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | 0.0599 | -0.101 | 0.0315 | -0.0819 | 0.0920 | 0.0357 | -0.0905 | 0.153 | 0.360** | 0.334*** | -0.169 |
| | (0.116) | (0.159) | (0.147) | (0.128) | (0.177) | (0.139) | (0.145) | (0.203) | (0.163) | (0.119) | (0.156) |
| DD [36-Month] Treatment* 36-Month | 0.304** | 0.416** | 0.638*** | 0.602*** | 0.580*** | 0.314* | 0.456** | | | -0.158 | 0.814*** |
| | (0.114) | (0.201) | (0.177) | (0.170) | (0.166) | (0.179) | (0.199) | | | (0.163) | (0.220) |
| DD [84-Month] Treatment* 84-Month | 0.0228 | -0.193 | -0.00492 | 0.198 | 0.0802 | -0.229 | -0.133 | -0.102 | -0.456** | -0.279 | -0.0650 |
| | (0.263) | (0.214) | (0.229) | (0.237) | (0.208) | (0.223) | (0.189) | (0.218) | (0.200) | (0.174) | (0.158) |
| Observations | 990 | 977 | 990 | 987 | 979 | 987 | 984 | 660 | 660 | 1,280 | 2,861 |
| R-squared | 0.111 | 0.068 | 0.212 | 0.168 | 0.148 | 0.147 | 0.054 | 0.076 | 0.038 | 0.075 | 0.108 |
| p-value (36m=84m) | 0.266 | 0.00200 | 0 | 0.00900 | 0.0220 | 0.00200 | 0.00100 | 0.518 | 0.0240 | 0.486 | 0 |

Notes: Estimations use difference in difference modelling using the balanced panel sampled and three waves (baseline, 36 months and 84 months). Income & Revenues and Finance & Debt indices are estimated using only the 36 and 84 month waves. Consumption, food security and child material needs indices are standardised measures against the control group within each round; the remaining summary indices are computed as the equally weighted average of z-scores of each indicator within the domain, then standardised against the control group within each round. Impact estimates are mean standardised ITTs, therefore effect sizes are expressed in SD of the control group. Robust standard errors clustered at the community level are in parentheses. *p < 0.1 **p < 0.05; ***p < 0.01. Estimations are adjusted and include respondent's age, education and marital status, household size and household demographic composition, and districts. Specifications 10 and 11 also include child's age and gender as controls.

### 4.8 Power and Multiple Inference Testing

Power: We provided sample size requirements in our pre-analysis plan to detect an effect of at least 0.20 SD for consumption, the livestock index, and several individual productive assets, focusing on indicators that had the largest sample size requirement. Our overall lead indicator of graduation out of poverty is consumption. For that indicator, and based on actual design effects from prior survey waves (intra-cluster correlations, number of clusters and number of households per cluster), we estimated a required sample size of 1,391 households, while for the livestock index our required sample size was just 239 households.

Our main results are in Table 9 on fade-out, estimated on SCT-ineligible households. Our balanced panel contains 1,615 households, so we have enough power to detect at least a 0.2SD effect. Row two of Table 9, which essentially replicates Handa et al (2018), shows all effect sizes to be greater than 0.20 and these are all statistically significant. In row three, all effect sizes that are less than 0.2SD are in fact not significant; the income and revenue effect side is 0.27 SD and is significant at 10 percent. The bottom row is our most directly relevant test, which is the difference between the 36-month and 84-month coefficients. In every case where the 36-month effect is statistically significant, the difference between the 36-month and 84-month coefficient is greater than 0.20 SD, and the associated p-values are less than 0.05 indicating statistical significance.

Power does become an issue when we move to Tables 12 and 13 that focus on the highly exposed sub-sample. In Table 12 we have just 881 observations, large enough for some of the productive indexes but not large enough for consumption. Indeed, there are a few cases where the 36-month and 84-month coefficients are larger than 0.20 SD but the p-value indicates non-significance (such as the productive asset index, the livestock index, and the food security scale). It is thus possible that some 36-month impacts are actually sustained in this sub-sample, but we do not have enough power to detect that.

Multiple Inference Testing: In our original 36-month analysis reported in Handa et al (2018), we accounted for multiple inferences by creating lead indicators or indexes, and by implementing the Sidak-Bonferroni adjustment. Here we have maintained the lead indicator/index approach but did not also provide adjusted p-values because virtually all the 84-month impact estimates are zero.

## 5. Challenges and Lessons

The main challenge in this study is the linking of programme data to our evaluations ample and accurately identifying HSCT eligibility status and CGP graduation dates. Ultimately these records had to be hand matched at the district field offices. As the HSCT is a new programme and targeting still underway in Western province, identifying current eligible households in our sample was manageable, though tedious. However, identifying accurate graduation dates for previous CGP households was fraught with difficulty, as the programme had ended several years ago in Kaputa), old records were not kept, and the transition of some study areas into new districts meant that many records were simply lost in transition. In other Transfer Project studies such as in Malawi and Ghana, strong efforts are made to facilitate linking of evaluation survey data to programme records, typically by sampling for a database provided by the Ministry and

maintaining a unique identifier across databases. This was not possible in the Zambian case because no systematic electronic records were kept at the district level at the beginning of the study.

The second challenge in the study was the remote locations of the study sites. Kaputa is a two-day journey by car from Lusaka, with the last half-day over treacherous, rocky terrain. Enumerators sleep in tents, households are isolated, and the field team carries all supplies, including petrol, with them. This makes field costs extremely expensive and physically challenging.

On the analysis side, the study team is still contemplating other ways to approach the analysis. One idea is to estimate the total sum of money received by a household and use that as the treatment effect. This would then also incorporate the ZMW 500 lump sum received by ineligibles in control areas. The team also intends to look in more detail at T-ineligible households that were able to maintain their consumption at 84-months, and see what their characteristics are, and identify what actions they took to maintain their consumption (e.g. invest in non-farm enterprise). The MCDSS is particularly interested in this analysis as it will help them identify complementary productive interventions for particular sets of households.

A final issue is that of reporting error in consumption and how this might influence the results. Specifically, under-reporting of consumption might be greater among T-ineligibles who might think that by reporting low consumption they could re-qualify for a cash grant. This incentive would not be as strong among C-ineligibles who received the lump-sum transfer and would not have an expectation of qualifying for a programme that they never had.

## 6. Discussion, Policy Implications and Conclusion

Our main finding is that large and across the board impacts of the CGP that occurred at 36-months, and which led to a sizeable multiplier effect, have disappeared after households left the programme. In other words, the original results, even though they encompassed economic investment and productive activity, were not sustained. This is driven mostly by fade-out but also some catch-up, and the pattern depends on the type of outcome we examine. Both consumption and food consumption suffer slight declines among the T-ineligibles and improve slightly among the C-ineligibles leading to convergence. On the other hand, asset and livestock indicators are maintained by the T-ineligibles in the face of growth among the C-ineligibles, which leads to some convergence. Spending on agricultural inputs and value of harvest is also maintained or increasing among T-ineligibles, but at faster rates among C-ineligibles resulting in convergence but with levels that are still higher in the original treatment group. The results on assets are similar to those reported recently by Haushofer and Shapiro (2018) for Give Directly. Perhaps the most telling result is that on subjective well-being, which shows a large drop among T-ineligibles but an increase in C-ineligibles. This could simply be a ‘hangover’ effect from suddenly being taken off the CGP rather than a real indication of material well-being.

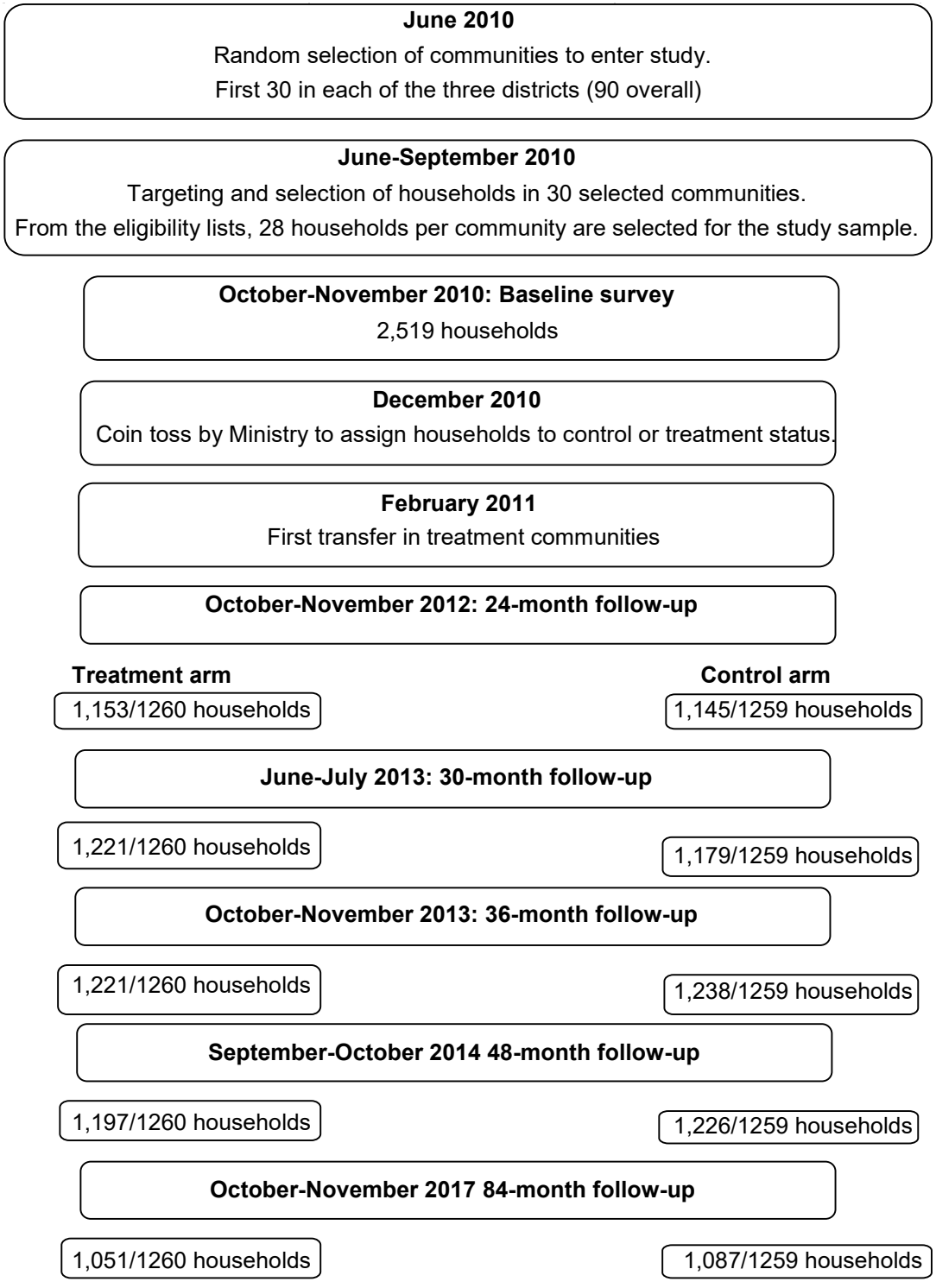
There is the possibility that under-reporting of consumption might be greater among T-ineligibles relative to C-ineligibles, which would bias the results towards fade-out.

However, the food security scale and the single indicator of ‘worried about food’ also decline, as do savings, suggesting that the decline in consumption and food consumption is real. On the plus side, asset levels and agricultural output is maintained or even increased.

There are several policy implications of these results. Households in this study are ultra-poor, with mean consumption of US$0.30 per person per day, some of the poorest households in the world. Infrastructure and environmental conditions are likewise quite harsh. The pattern of results we find, that large programme effects mostly fade-out quickly, suggest that in this environment, and with households at the edge of survival, cash alone is unlikely to lead to wholesale graduation out of poverty. For the HSCT specifically, current programme rules are that HSCT recipients are enrolled for three years, after which they must undergo a recertification process to keep their eligibility. Our results indicate that even after three years of cash transfers, consumption, food security and other protection outcomes are likely to decline if the households are removed from the programme, especially HSCT beneficiaries who tend to be less economically viable due to the demographic eligibility criterion. An immediate implication is that the re-certification process might be extended up to five years to save resources.

In terms of future work, the next step is to identify households that were able to maintain or even increase their consumption, and see what actions they took in previous years to enable them to maintain their trajectory. This in turn will provide insight on complementary interventions to help households graduate from poverty.

## Appendix A and B: Study Flow Chart and Timeline



## Appendix C: Map of Zambia and Study Districts

## Online appendix D: Indicator definitions

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-D-Indicator-definitions.pdf

## Online appendix E: Additional Attrition Results – Overall Attrition

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-E-Additional-Attrition-Results%E2%80%93Overall-Attrition_0.pdf

## Online appendix F: Differential Attrition SCT Ineligible Group

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-F-Differential-Attrition-SCT-Ineligible-Group.pdf

## Online appendix G: Differential Attrition SCT Eligible Group

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-G-Differential-Attrition-SCT-Eligible-Group.pdf

## Online appendix H: Tables of means and simple difference in difference by indicator (in actual units) and treatment status, among HSCT eligible and ineligibles

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-H-Tables-of-means-and-simple-difference.pdf

## Online appendix I: Further impact analysis and/or robustness checks

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-I-Further-impact-analysis.pdf

## Online appendix J: Pre-Analysis Plan

https://www.3ieimpact.org/sites/default/files/2019-08/DPW1.1042-Zambia-UCT-Online-appendix-J-Pre-Analysis-Plan.pdf

## References


AIR 2011. Zambia's Child Grant Program: Baseline Report. Washington DC.

AIR 2014. Zambia's Child Grant Program: 36-month impact report. Washington, DC: Author.

Baird, S., McIntosh, C. and Özler, B., 2016. When the money runs out: do cash transfers have sustained effects on human capital accumulation?. The World Bank.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B. and Udry, C., 2015. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. Science, 348(6236), p.1260799.

Blattman, C., Fiala, N. and Martinez, S., 2018. The long term impacts of grants on poverty: 9-year evidence from Uganda's Youth Opportunities Program (No. w24999). National Bureau of Economic Research.

Blattman, C. and Niehaus, P., 2014. Show them the money: Why giving cash helps alleviate poverty. Foreign Affairs, 93(3), pp.117-126.

Brown, C., Ravallion, M. and Van de Walle, D., 2016. A poor means test? Econometric targeting in Africa. The World Bank. Policy Research Working Paper No.7915.

Edmonds, E.V. and Shrestha, M., 2014. You get what you pay for: Schooling incentives and child labor. Journal of Development Economics, 111, pp.196-211.

Garcia, M., Moore, C., 2012. The Cash Dividend: The Rise of Cash Transfer Programs in Sub-Saharan Africa. Washington, DC, The World Bank.

Handa, S., Natali, L., Seidenfeld, D., Tembo, G., Davis, B. and Zambia Cash Transfer Evaluation Study Team, 2018. Can unconditional cash transfers raise long-term living standards? Evidence from Zambia. *Journal of Development Economics*, 133, pp.42-65.

Handa, S., Seidenfeld, D., Davis, B., Tembo, G., the Zambia Cash Transfer Evaluation Team. 2016a. The Social and Productive Impacts of Zambia's Child Grant. *Journal of Policy Analysis and Management*. Volume 35, Issue 2, pages 357–387, Spring 2016

Handa, S., Natali, L., Seidenfeld, D. and Tembo, G., 2016. The impact of Zambia's unconditional child grant on schooling and work: results from a large-scale social experiment. Journal of Development Effectiveness, 8(3), pp.346-367.

Haushofer, J., Shapiro, J. 2016. The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *Quarterly Journal of Economics*, 131 (4): 1973-2042.

Haushofer, J. and J. Shapiro. 2018. The long-term impact of unconditional cash transfers: Experimental evidence from Kenya. Working Paper, Princeton University.

Stoeffler, Q., Mills, B. and Premand, P., 2016. Poor households' productive investments of cash transfers: quasi-experimental evidence from Niger. Policy Research Working Paper Series 7839, The World Bank.

University of North Carolina, 2016, Malawi Social Cash Transfer Program: Endline Impact Evaluation Report, Carolina Population Center.

University of North Carolina, 2018, Zimbabwe Harmonized Social Cash Transfer Programme: Endline Impact Evaluation Report. Available from Carolina Population Center and UNICEF Zimbabwe.

World Bank 2015. The State of Social Safety Nets 2015. Washington, DC: World Bank.