World Development 175 (2024) 106483

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Regular Research Article

Unconditional cash transfers and safe transitions to adulthood in Malawi

Monica Lambon-Quayefio [a], Amber Peterman [b,*], Sudhanshu Handa [b], Adria Molotsky [c], Frank Otchere [d], Peter Mvula [e], Maxton Tsoka [e], Jacobus de Hoop [f], Gustavo Angeles [b,g], Kelly Kilburn [h], Annamaria Milazzo [i], on behalf of the Malawi Social Cash Transfer Evaluation Team

[a] Department of Economics, University of Ghana, Accra, Ghana
[b] Department of Public Policy, Abernathy Hall, 131 S. Columbia S. CB #3435, Chapel Hill, NC 27599-3435, United States
[c] American Institutes for Research (AIR), Washington D.C., United States
[d] Social and Economic Policy Unit, UNICEF Office of Research—Innocenti, Florence, Italy
[e] University of Malawi, Zomba, Malawi
[f] World Bank, Washington D.C., United States
[g] Department of Maternal and Child Health, University of North Carolina at Chapel Hill, Chapel Hill, United States
[h] Duke Global Health Institute, Duke University, Durham, NC, United States
[i] Institute for Research on Women, Gender, and Sexuality (IRWGS), University of Houston, Houston, United States

ARTICLE INFO

Keywords:
Cash transfers
Safe transitions
Malawi
Gender
Impact evaluation

ABSTRACT

As national social safety nets are expanding in Africa and globally, evidence on the impact of programs on youth transitions can help guide future investment and program design decisions. This paper examines the effects of Malawi's flagship cash transfer program on safe transitions to adulthood among youth living in households experiencing ultra-poverty. The evaluation was a cluster-randomized control trial implemented over three years using panel data on youth aged 13 to 19 at baseline. Household receipt of bi-monthly transfers led to improvements in four out of six outcome domains: 1) education, 2) physical health, 3) emotional wellbeing and mental health, and 4) sexual and reproductive health. There were limited or no impacts on outcomes related to HIV risk and time use. Results imply that similar programs in the region may also facilitate safe youth transitions. Nonetheless, more intentional gender- and youth-specific designs may have promise for more holistic impacts, and further evidence is needed on longer-term effects.

1. Introduction

Adolescence is increasingly recognized as a key window of developmental opportunity, with investments yielding 'triple dividends' on the current health and wellbeing of adolescents, their future trajectories into adulthood, and on the health and wellbeing of their children (Dahl et al., 2018). Physical maturation, brain development and social changes coincide with transitions starting in adolescence, which set the trajectories and opportunities for health and wellbeing into adulthood. In Africa, the population of adolescents and youth aged 15 to 24 is projected to grow faster than in any other region, with an estimated near doubling between 2015 and 2050 (from 258 million to 456 million) (United Nations, 2020). For this reason, adolescents and youth have been recognized as a key demographic group for the region, for harnessing the economic and social benefits of the demographic dividend, and critical for achievement of the Sustainable Development Goals (SDGs). With many countries in the region lagging in their progress towards the SDGs, investment in policies and programming geared towards the improvement of adolescent and youth health and wellbeing is needed.

Social safety nets have been widely used to protect the most marginalized and those who experience extreme forms of poverty in the face of shocks and facilitate investment in health and education. In Africa, on average, countries spend approximately 1.2 % of GDP or about 4.6 % of total government spending on social safety net programs (Beegle et al., 2018). Despite abundant evidence on their beneficial

* Corresponding author.
E-mail addresses: mplambon-quayefio@ug.edu.gh (M. Lambon-Quayefio), amberpeterman@gmail.com (A. Peterman), shanda@email.unc.edu (S. Handa), amolotsky@air.org (A. Molotsky), fotchere@unicef.org (F. Otchere), petermvula58@yahoo.com (P. Mvula), tsokamax@gmail.com (M. Tsoka), jdehoop@worldbank.org (J. de Hoop), gustavo_angeles@unc.edu (G. Angeles), kilburn.kelly@gmail.com (K. Kilburn), amilazzo@uh.edu (A. Milazzo).

https://doi.org/10.1016/j.worlddev.2023.106483
Accepted 22 November 2023
Available online 2 December 2023
0305-750X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

impacts, there is little research on the ability of these programs to facilitate inter-generational impacts on adolescent and youth safe transitions.[1] Available evidence is either concentrated in Latin America, focusing on inter-generational impacts on education and labor force participation (e.g., Millán et al., 2019), or among adolescent-focused smaller programs implemented by NGOs, often linked to the HIV sector in Africa (e.g., Pettifor et al., 2016; Baird et al., 2012). There are numerous ways social safety nets and cash transfers might influence safe transitions, including through household wealth and education effects, as well as potential for reductions in risky behaviors, including transactional sex, early marriage and child labor (Cirillo et al., 2021; Handa et al., 2015; Heinrich et al., 2017). However, these effects are not automatic, particularly when economic support might be targeted at the household-level and used primarily for household consumption needs. Understanding if large-scale cash transfers have positive outcomes for youth matters for future investment in cash transfer programs. This holds especially for unconditional programs, not tied to a specific investment behavior, which have the potential to change wellbeing outcomes across a broad range of multi-dimensional domains.

This paper examines the effects of the government of Malawi's flagship Social Cash Transfer Program (SCTP, locally known as *Mtukula Pakhomo*) on safe transitions to adulthood among adolescents living in households experiencing ultra-poverty. Transfers were given unconditionally to heads of households, valued at approximately 17 to 23 % of median pre-program household consumption. The program was evaluated using a cluster randomized control trial (RCT) corresponding to the government roll-out of the program, during the period 2013 to 2015—consisting of three waves approximately 12 months apart. We use Analysis of Covariance (ANCOVA) models to analyze indices across six outcome domains collected among youth aged 13 to 19 at baseline: 1) physical health, 2) emotional wellbeing and mental health, 3) education, 4) time use, 5) sexual and reproductive health, and 6) HIV risk. A total of 25 indicators are represented across the six domains. Further, we examine if impacts accrue over time (from midline to endline) and if impacts differ between male and female youth. As the program is unconditional and cash transfers are given to heads of households, we hypothesize that the program will work via economic channels, facilitating increased education, better home environment, health and reduced need to engage in risky behaviors for youth.

Our findings indicate that household receipt of bi-monthly transfers improved youth outcomes on average in all but one domain examined (time use). Improvements range from 0.09 to 0.30 standard deviations (SDs) in comparison to the control group and accrue both in domains thought to be more directly connected to income (e.g., education), as well as more distal or complex domains (e.g., sexual and reproductive health). For domains of physical health as well as emotional wellbeing and mental health, impacts accumulate over time, while others remain constant across study waves. For one domain, HIV risk, impacts are only significant at midline, and in less preferred model specifications, thus can be interpreted as minimal. In only one domain (physical health), impacts are significantly larger among female youth, suggesting the program generally benefited both sexes to equal degree. The lack of beneficial impacts on time youth spend in domestic and productive work may be in part due to increases in non-farm enterprises and other income generating activities at the household level. Thus, while there were large impacts on enrollment and spending on education—these changes in youth activities did not translate to less time spent in productive work or domestic chores, offset by increases in productivity at the household level.

This paper contributes to a growing literature on the impacts of social protection on adolescent safe transitions. A review of government non-contributory programs in low- and middle-income countries (LMICs) finds 85 studies corresponding to 34 different programs, however most of the literature is focused on education and child labor impacts (Cirillo et al., 2021). More diverse outcomes, including mental health, sexual and reproductive behavior and HIV risk are limited to a handful of studies each, although a broader range of studies exist examining NGO-implemented programming (de Walque et al., 2017; Stoner et al., 2021). Our paper is most similar to three studies of cash transfers examining adolescent health and wellbeing in southern Africa. The first study examines the impacts of an HIV Prevention Trials Network conditional cash transfer, on multi-dimensional deprivation of female youth of secondary school age over a four-year period from 2011 to 2015 in South Africa (Kilburn et al., 2020). Using an experimental design, the authors find protective impacts of the program on overall deprivation, driven by three domains: protection, relationships and economic agency. The second study of an NGO-implemented cash transfer in Malawi examined multi-dimensional impacts of two years of transfers on female youth aged 13 to 22 and unmarried at baseline. Results show initial impacts on multiple domains, including education, HIV and sexually transmitted infections, psychological distress, pregnancy and early marriage, with differences in impacts depending on whether transfers were given unconditionally or conditional on school enrollment (Baird et al., 2011, 2012). However, at a two-year post-program follow-up, many of these impacts had dissipated, especially in the unconditional arm—where there were no initial impacts on schooling outcomes (Baird et al., 2019). The third study examines the relationship between South Africa's Child Support Grant, a government unconditional cash transfer on adolescent risky behavior (Heinrich et al., 2017).[2] The study finds the dosage (i.e., number of months) adolescents reside in households receiving the grant is linked to delayed sexual debut, pregnancy and number of sexual partners for females, and reduced probability of drinking for males—however it does not examine broader domains of adolescent wellbeing or identify causal impacts.

Our paper differs from these studies in several ways. First, this study includes a wide range of outcomes and examines impacts on both male and female youth. Thus, we are able to assess if young men benefit from the program in a holistic way as compared to young women. In addition, unlike the aforementioned studies in Malawi and South Africa (Baird et al., 2019; Kilburn et al., 2020), the current study examines a government program operating at scale in rural areas and has common features to national programs in the region. Like many other programs focused on labor-constrained households, the Malawi SCTP was not explicitly designed to improve transitions to adulthood. However, since school enrollment is one of several program objectives, and since transfer values are larger in households with higher numbers of secondary-school aged youth, the program has some implicit design elements with a focus on youth. The vast majority of evaluation evidence on cash transfers are still concentrated in Latin America, where context, design and safe transitions occur at different rates (Millán et al., 2019). In addition, most evidence on cash transfer programs and safe transitions examines specific domains or indicators, rather than taking a holistic view on impacts (Cirillo et al., 2021; de Walque et al., 2017). These results serve as a basis for further discussions on intentional program design components to facilitate safe transitions to adulthood for youth in the context of national programs in the Africa region.

Our results add to broader evidence from the Malawi SCTP showing the program was successful at reaching its main objectives by decreasing household poverty, improving consumption, food security, asset ownership and resilience (Handa et al., 2022; Otchere & Handa, 2022).

[1] While there may be different interpretations of what 'inter-generational' means in the context of social programs, we define it to mean impacts on the current or future children of primary participants (target individuals) (i.e., the next generation) (Parker & Vogl, 2018).

[2] The study uses a partial identification model and the staged expansion age of eligibility within the Child Support Grant, thus interpreting results as associations linked to larger doses (durations) of the cash transfer, however, is careful to add caveats that they are not able to identify causality.

In addition, analysis of school-aged children 6 to 15 years shows increases in school enrollment, regular attendance, schooling expenditure and material wellbeing (de Hoop et al., 2020). However, this analysis also indicates that work increased among school-aged children as children may have been both drawn into entrepreneurial activities and labor substitution dynamics took place in the household. Finally, three distinct outcomes have been examined related to safe transitions to adulthood among adolescents and youth. Examination of all youth living in households aged 14 to 21 at baseline showed no impacts on early marriage for both males and females, or early pregnancy and childbirth for females (Dake et al., 2018). Further, improvements in mental health and decreases in depressive symptoms, as measured by the Center for Epidemiological Studies Depression (CES-D) scale were found among adolescents aged 13 to 19 (Angeles et al., 2019). The latter study is the only publication to date using the youth module data and which overlaps explicitly with the analysis in the current paper. Angeles and colleagues (2019) examine impacts on small fraction of outcomes included in this analysis, namely mental health and pathways at endline (mental health only for the male sample, and mental health plus five additional pathway indicators for the female sample).[3] We complement previous analyses by examining the scope of holistic impacts, including additional domains, indicators and measuring impacts across waves.

2. Malawi's social cash transfer program

The Malawi SCTP is an unconditional cash transfer program targeted to ultra-poor, labor-constrained households, administered by the Ministry of Gender, Children, Disability and Social Welfare with policy coordination and oversight by the Ministry of Finance, Economic Planning and Development.[4] The overarching objective of the SCTP is to reduce extreme poverty and hunger, and to increase school enrollment among households experiencing ultra-poverty. The program began as a pilot in the Mchinji district in 2006, was expanded to an additional eight districts by 2009, and further expanded starting in 2014. As of 2020, the program reached approximately 1.26 million individuals in 275,000 households in all 28 districts (or approximately 6.5 % of the population).[5] During the evaluation period, participant households received bi-monthly payments of varying amounts depending on household size and the number of primary and secondary school-aged children in the home. At the time of the endline survey for this study, the average monthly per capita transfer amount received by participating households was equivalent to US$1.25 a month or about US$60 per household annually. Payments were made in cash every other month through a local pay-point manager to the main beneficiary, and there were no conditions to receive the money. Payments represented approximately 17 to 23 % of median pre-program household consumption for eligible households. In May 2015, due to inflation and decline in the real value of transfers, payment amounts were increased. A summary of the benefit levels by household characteristic and by year is provided in Appendix Table A1. Although not a condition or explicit target group of the program, because of the focus on ultra-poor and labor constrained households, the majority of adult recipients are female (84 % of the evaluation sample) (Carolina Population Center, 2014, 2016).

3. Data and methodology

3.1. Data

The scale-up of the SCTP was evaluated through the Transfer Project, a research and learning consortium focused on government cash transfers in Africa, led by the University of North Carolina, in partnership with the UNICEF Office of Research—Innocenti and the University of Malawi. This evaluation is a cluster RCT implemented during the expansion of the government's program in rural areas of Salima and Mangochi districts in central and southern regions of Malawi, respectively (see Fig. A1).[6] The sample selection occurred as follows: First, two traditional authorities (TAs) in each district were randomly selected for inclusion in the evaluation. Second, village clusters in each TA were randomly ordered. Third, eligible households within each village cluster were interviewed starting with the first ranked village cluster and continuing until the required sample size was met. This resulted in a total of 29 village clusters included in the study. The randomization of village clusters to the study arms was implemented via public lottery with the Salima and Mangochi district commissioners' offices and stratified by TA resulting in 14 treatment and 15 control village clusters. The randomization occurred after the baseline survey to 'blind' the treatment and control groups to the survey teams and maintain internal validity of the study. Overall, there were three surveys waves between 2013 and 2015: a baseline, midline (17-month follow-up) and endline (30-month follow-up). At midline, eligible households had received between five to six bi-monthly payments from the SCTP. At endline, eligible households had received a total of 12 bi-monthly payments. Data collection was led by the Center for Social Research at the University of Malawi, conducted in local languages, *Chichewa* or *Chiyao,* and captured using electronic tablets. For all three survey rounds, experienced enumerators received extensive training on survey instruments and the ethics of human research to ensure adherence to high research ethics. The study underwent ethical clearance by the University of North Carolina Internal Review Board (IRB), Malawi's Commission for Science and Technology and Malawi's National Committee for Research in Social Sciences and Humanities.

There were six main research questions agreed upon by stakeholders for the overall SCTP evaluation (e.g., does the SCTP improve consumption, reduce food insecurity, increase dietary diversity?; does the SCTP affect economic productivity and wealth accumulation?; does the SCTP affect health and nutrition of young children?) (Carolina Population Center, 2014). One of these related specifically to safe transitions to

[3] Specifically, Angeles and colleagues (2019) include impacts on: The Center for Epidemiological studies depression (CES-D) scale (for male and female youth), as well as school enrollment, grade attainment, reported health, incidence of mobility and social support (for female youth only). The previous analysis includes endline impacts only, thus reports on 2 percent (1 out of 50) of the impacts for male youth and 12 percent (6 out of 50) of the impacts for female youth included in the current analysis.

[4] The Operations Manual of the SCTP defines ultra-poor as a household that is unable to meet the most basic urgent needs, including food and essential non-food items such as soap and clothing, while labor constrained refers to a household with a dependency ratio (ratio of 'not fit to work' to 'fit to work') of more than three or with no individuals who are fit to work. An individual who is considered not fit to work is someone who meets one of the following conditions: 1) under the age of 18; 2) over the age of 64; 3) within the age range 18 to 64 but suffering from a chronic illness or disability or is otherwise unable to work (Carolina Population Center, 2014).

[5] Participant figures come from the SCTP website and use the UN 2020 mid-year population estimates to calculate population coverage: https://mtukula.com/index.php.

[6] While no comparative poverty statistics exist from the baseline data collection period by district, our sample population can generally be thought of as similar to the rural population of ultra-poor in other central and southern regions. Overall the rates of ultra-poverty in Malawi in 2013 were 11.5% nationally and 12.3% in rural areas, with 91.1% of all ultra-poor residing in rural areas—thus the evaluation sample represents the very poorest segment of the population in each district (National Statistics Office, 2014).

adulthood among adolescents, which was identified as an under-researched area at the time of study design. Therefore, in all waves, in addition to interviewing the head of household, separate youth survey modules were administered by enumerators of the same sex to adolescents aged 13 to 19 years old at baseline. A maximum of three individuals within this age bracket were interviewed per household, due to budgetary and logistical constraints. The youngest three adolescents were prioritized in instances where a household had more than three individuals who fell within the age bracket. However, this rule was not strictly followed, for example, exceptions were made if there were logistical considerations favoring an older adolescent due to his/her availability at the time of the interview. These interviews were conducted in private due to the sensitive nature of the questions and enumerators were guided on strategies to ensure privacy and confidentiality. Informed consent for the youth interview was obtained either from parents of youth less than 18 years (and assent obtained from the youth themselves) or obtained directly from youth who were 18 years and above. The sample size for the study was originally set at 3,300 households, powered based on the indicator for which the study team believed would be most challenging to change (child nutritional status).[7]

Questions specific to the youth module include schooling, mental health, future aspirations and expectations for future quality of life, time preferences, personal and extended family social networks, sexual activity and risk-taking behaviors. The youth module was designed based on existing surveys and evidence at the time of the study (approximately 2013–2015) around potential impacts of UCTs in the region. We use this rich data on multiple domains representing youths' safe transitions, focusing on the panel of youth who were interviewed at baseline and again at one or both follow-up waves. We complement data from the youth survey with standard data on demographics and background characteristics from the household survey. Additional details on the data collection and survey implementation can be found in the endline technical report (Carolina Population Center, 2016).

3.2. Outcome indicators

Existing evidence indicates cash transfer interventions can impact a wide range of outcomes for households and individuals, through poverty and human capital formation channels, among others. We draw on an adapted version of the Gender and Adolescent Global Evidence (GAGE) conceptual framework for adolescent wellbeing and development to inform our analysis (GAGE Consortium, 2017). The framework uses 'capabilities, change strategies and context' to understand what works to support adolescent girls focusing on six main domains. Following this framework, we analyze outcome indices across six domains relevant for safe youth transitions: 1) physical health, 2) emotional wellbeing and mental health, 3) education, 4) time use, 5) sexual and reproductive health, and 6) HIV risk. Our domains differ somewhat from the GAGE conceptual framework, guided by differences in context and hypothesized theory of change for a cash transfer intervention specifically (which affected the inclusion of questions in the household and youth surveys).[8] For each domain, we consider all relevant outcomes represented in the youth module, and exclude only those with very low prevalence due to power issues (low prevalence also may signify lower relevance to the study population in capturing safe transitions).[9] Each resulting domain is made up of three to six individual indicators, representing 25 indicators in total. For example, the physical health domain is comprised of four indicators: 1) self-reported health rated on a one to five point scale (ranging from poor to excellent), 2) free from illness or injury in the last two weeks (binary indicator), 3) sum of responses for limitations in five activities of daily living (ADL) tasks (seeing, hearing, walking / climbing steps, remembering or concentrating, and communicating), each ranging from zero to three, and 4) if the youth slept under a mosquito net last night (binary indicator). Appendix Table A2 lists all the indicators and definitions used in the construction of these indices for each domain. It should be noted that some indicators were only collected at one follow-up wave, for example, time use indicators were only collected at endline (wave 3) and not midline (wave 2). For other indicators, such as aspirations and social support, we lack baseline data for the measures, thus estimations do not control for baseline values.

To create the aggregate index, each indicator is standardized as a z-score in relation to the control mean by wave, aggregated and re-standardized (Kling et al., 2007). For ease of interpretation, for all indicators, scales are transformed so that higher values correspond to better wellbeing or a favorable transition. In the construction of indices, we replace missing values for individual indicators with averages by wave and treatment status to maintain consistent sample sizes across domains. In practice, the absolute number and percentage of missing values for individual indicators is extremely low (ranging from one to 34 observations, in all cases less than 1 % of the total sample).

3.3. Modeling approach

The estimation follows a standard ANCOVA approach, regressing the outcome at follow-up wave(s) on the treatment indicator, a set of control variables, and the outcome at baseline. ANCOVA is preferred to difference-in-difference (DD) models due to power gains when autocorrelations between baseline and follow-up dependent variables are low. Guidelines proposed in (McKenzie, 2012) suggest "low"

[7] Power calculations used data from the earlier Mchinji pilot study of the SCTP based on children aged 0 to 5 years in target households. Part of the rationale for choosing child nutrition was the literature indicating challenges in changing nutritional status without holistic intervention. In addition, since the SCTP targets labor-constrained households, it was expected that not all households would have a child in this age range. The power calculations showed that a change of 0.225 standard deviations in height-for-age with an alpha level of 0.05 and 80 percent power (accounting for design effects and 10 percent attrition rates) would be approximately 3,300 households split evenly between treatment and control clusters (Handa et al., 2013).

[8] The domains proposed in the GAGE Conceptual Framework are: 1) bodily integrity and freedom from violence, 2) psychosocial well-being, 3) education and learning, 4) economic empowerment, 5) sexual and reproductive health and nutrition, 6) voice and agency. We follow these groupings, however for the first domain, we include physical health more broadly, rather than bodily integrity, specifically. In addition, we did not collect indicators related to domain 6 (voice and agency), however create a separate domain HIV risk, distinct from other sexual and reproductive health measures, as this is an area of high importance and emphasis in Malawi. The original conceptual framework for the SCTP proposes five similar domains for older children as possible outcomes (mental health, schooling, material wellbeing, work, and HIV risk). Full youth surveys across waves for this work can be found on the Transfer Project webpage: https://transfer.cpc.unc.edu/countries/malawi/#instruments.

[9] For example, we exclude alcohol and tobacco use, as prevalence levels for lifetime use are very low at baseline (<2% of the sample) and unlikely to show sensitivity to the intervention. Likewise, a time preference measure (willingness to wait for monetary payouts) was collected at endline, however there was very low variation in this outcome, signaling that youth did not understand questions well, or payout amounts were not variable enough to capture nuances in time preferences. Although no pre-analysis plan was undertaken for this analysis, all other indicators were defined and discussed among the author team before analysis.

autocorrelation is between 0.20 and 0.40, while "high" autocorrelation is between 0.60 and 0.80. As shown in Appendix Table A3, across 17 individual outcomes that were collected both at baseline and at follow-up, only one (highest grade of schooling completed, correlation: 0.77) can be considered "high" and all but one domain index (education) can be considered "low." This is expected, as indicators like 'highest grade completed' have high path dependence over data collection waves because youth who leave school are unlikely to re-enroll over time. Thus, we consider ANCOVA to be a preferred estimator and examine the robustness of our findings to DD modeling techniques.

The overall program impact is estimated based on the following specification using a pooled sample of both midline and endline follow-ups:

$$Y_{ihj1} = \beta_0 + \beta_1 Treat_j + \beta_2 Y_{ihj0} + X_{ihj0} + Endline_{ihj1} + \gamma_{TA} + \varepsilon_{ihj1} \quad (1)$$

Where Y_{ihj1} represents the domain of interest observed at endline for youth i in household h and village j. Similarly, Y_{ihj0} indicates the domain of interest observed at baseline. The binary indicator $Treat_j$ is used to capture the treatment status of each village, taking on a value of 1 if village j receives the cash transfer and 0 otherwise. X_{ihj0} shows baseline covariates, and $Endline_{ihj1}$ is a binary variable equal to 1 if the follow-up is the endline (where 0 is the midline). These covariates include an indicator for sex (female) and age of youth (in year fixed effects), an indicator for sex (female) and age of household head (in continuous years), an indicator for whether the household head ever attended school, as well as household size. The covariates were chosen based on key socio-economic and demographic characteristics likely to be important in determining key youth wellbeing outcomes, following previous analysis of youth and household-level outcomes (Angeles et al., 2019; Handa et al., 2022). We also include district-level TA indicators—the level of stratification at randomization—represented by γ_{TA} to capture any differences in program impacts at the TA level. The coefficient β_1 captures the intent-to-treat impact of the intervention on each domain in equation (1), controlling for baseline values, time trend and covariates. All standard errors are clustered at the village level. Further, we implement the wild bootstrap with 1,000 replications to account for the fact that we only have 29 clusters (Cameron et al., 2008; Roodman et al., 2019).

We consider two types of heterogeneous impacts. First, impacts are differentiated by sex. Second, we explore the evolution of program impacts over time. For each heterogeneous impacts under the ANCOVA specification, we include female and wave interaction effects in the base specifications as shown in equations (2) and (3) respectively.

$$Y_{ihj1} = \beta_0 + \beta_1 Treat_j{*}Female_{ihj} + \beta_2 Treat_j + Female_{ihj} + Endline_{ihj1} + \beta_3 Y_{ihj0} + X_{ihj0} + \gamma_{TA} + \varepsilon_{ihj1} \quad (2)$$

$$Y_{ihj1} = \beta_0 + \beta_1 Treat_j{*}Endline_{ihj1} + \beta_2 Treat_j + Endline_{ihj1} + \beta_3 Y_{ihj0} + X_{ihj0} + \gamma_{TA} + \varepsilon_{ihj1} \quad (3)$$

Where $Female_{ihj}$ takes on a value of 1 if the individual is female and 0 if the individual is male. X_{ihj0} remains the vector of covariates as explained earlier (excluding sex of the individual in the case of equation (2)). In addition to models (2) and (3), for illustrative purposes, we estimate equation (1) separately for the male and female samples (midline and endline samples).

3.4. Sample and internal validity

Our analytic sample is based on 1,679 youth who were interviewed at baseline and again at least one follow-up wave (815 youth across 628 households in the treatment group and 864 youth in 648 households in the control group). Therefore, our panel is unbalanced, consisting of some youth interviewed in all three waves, and some interviewed in the baseline and just one of either midline or endline waves. Fig. A2 shows the evolution of the study sample, including details around the original selection of study area, selection of eligible households within these areas, youth interviewed within eligible households at baseline, and subsequent randomization and reinterview of youth at midline and endline survey waves.

Table 1 shows the background characteristics and aggregate outcomes of youth in the sample at baseline—including tests of balance between treatment and control groups. Youth are on average 15.3 years old, with modal years of 14 and 15 years (approximately 22.9 and 19.6 % of the sample, respectively) and female youth make up 47.4 % of the sample. Youth live primarily in female headed households (86 % of the sample), and heads of households are on average 53 years old. Average household size is 5.8 members. Across all background characteristics and five outcome aggregates collected at baseline, there is good balance, except for age 16 (the control group has a higher proportion of youth aged 16, p = 0.036) and the sexual and reproductive health score, in which the treatment is significantly worse off at baseline (p = 0.060). Table A4 further reports baseline balance for the individual index components. These results show that the imbalance in the sexual and reproductive health score whereby the treatment sample is worse off at baseline originates from two indicators: the partnership indicator (currently single – not married and no boyfriend or girlfriend, p = 0.093)

Table 1
Baseline balance for background characteristics and outcomes for youth.

Background characteristics	All	Control	Treatment	P-value from difference
Age 13 years	0.178	0.167	0.190	0.273
Age 14 years	0.229	0.231	0.227	0.843
Age 15 years	0.196	0.189	0.204	0.450
Age 16 years	0.132	0.147	0.117	0.036
Age 17 years	0.122	0.132	0.112	0.126
Age 18 years	0.095	0.091	0.098	0.678
Age 19 years	0.048	0.043	0.053	0.341
Youth is female	0.474	0.497	0.450	0.159
Household head is female	0.861	0.858	0.864	0.838
Household head age (years)	53.40	51.73	55.17	0.121
Household head ever attended school	0.381	0.392	0.369	0.745
Household size	5.835	5.927	5.737	0.386
District TA 1	0.310	0.352	0.266	0.656
District TA 2	0.273	0.247	0.301	0.770
District TA 3	0.179	0.167	0.191	0.861
Aggregate z-score outcomes				
Physical health scale	−0.016	−0.000	−0.032	0.779
Emotional wellbeing & mental health scale	0.040	−0.000	0.083	0.242
Education scale	−0.052	−0.000	−0.108	0.284
Sexual & reproductive health scale	−0.078	−0.000	−0.161	0.060
HIV risk scale	−0.062	−0.000	−0.127	0.106
N	1,679	864	815	

Notes: P-values are reported from Wald tests on the equality of means of control and treatment for each variable. Standard errors are clustered at the village cluster level. The time use domain was only collected at endline (wave 3) and thus it is not included.

and never been forced, tricked or pressured into sex (p = 0.066). While all models control for initial baseline levels, this indicates that the impacts on the aggregate score and individual indicators in this domain should be interpreted with some caution. Tables A5 and A6 replicate the main descriptive table showing background characteristics and aggregate outcomes for female and male youth, respectively. These results show similarly well-balanced background characteristics, however the imbalance in sexual and reproductive health carries over in the female youth sample (p = 0.063). In addition, there are imbalances in the HIV risk index for female youth (p = 0.012), suggesting some caution when interpreting the female-specific impacts on these domains.

We present detailed analysis of sample selection and attrition and show that these factors are unlikely to pose threats for internal validity. Since up to three eligible youth per household were selected to complete the youth module, we may be concerned that youth who are interviewed may be systematically different from youth who were not interviewed. This could occur because younger youth were prioritized, or because youth who were out of school or at home (not working) may be more likely to be available for the interview during the survey team visit. We investigate this possibility using household data comparing basic demographics and control variables collected as part of the household survey for interviewed youth (N = 2,109, approximately 76 %) versus other youth in the household of the same age range at baseline (N = 673). Table A7 shows that while youth selected for interview are significantly younger than youth not-selected, and (in treatment clusters) live in households with older heads, no other significant differences are detected. In addition, there are no significant differences between those not interviewed between treatment and control households (column 8), indicating the sampling strategy did not compromise internal validity.[10]

Attrition analyses show how sample attrition over survey waves may affect the internal validity of our findings. In particular, we may be concerned about systematic differences between treatment and control youth who are lost to follow-up. Overall, we find that of the 2,109 youth interviewed at baseline, 20.4 % were lost to follow-up by endline. These percentages are similar for the treatment and control (22.0 % versus 18.8 %) and not statistically significant at conventional levels (Table A8). We also examine background characteristics and aggregates (Table A9), as well as baseline outcome indicators (Table A10) to test if there are mean differences between groups lost to follow-up between treatment and control. The key results for these tests are found in column (8) which shows the p-value comparison between control attritors (in column 1) and treatment attritors (in column 4). Table A9 shows that, across 20 background characteristics and outcome aggregates, only two age indicators show signs of differential attrition—youth aged 17 and aged 19 are more likely to leave the control group (p = 0.068 and p = 0.096, respectively). In addition, only two out of 17 index components show signs of differential attrition (Table A10). Control youth who leave the sample have higher scores on ADLs (p = 0.035) but are more likely to have ever given or received money for sex with their most recent partner (p = 0.024). Tables A11 and A12 replicate this attrition analysis for the background characteristics and aggregate outcomes for female and male youth, respectively. Sub-sample results similarly show that differential attrition is unlikely to be a threat to internal validity. Similar to the full sample, female youth who are aged 19 years are more likely to leave the control group. In addition, male youth who leave the control group have higher aggregate physical health as compared to the treatment group leavers. For all other background characteristics and aggregate outcomes—attrition appears balanced between the treatment and control group. These analyses indicate that differential attrition is unlikely to be a large concern for the internal validity of our findings.

Finally, we may also be concerned that using the unbalanced panel (intersection of baseline and any follow-up) might be a concern—rather than just youth appearing in all three waves. To address this issue, we rerun our analysis using only full panel youth (N = 2,072) and find coefficients and significance levels are nearly identical to those in the full sample. However, in few cases, balanced sample estimates are slightly larger as compared to our unbalanced estimates (thus, we view our main estimates as conservative). These results indicate selective attrition across waves is unlikely to result in meaningfully different conclusions.

4. Results

4.1. Impacts on aggregate domains

Table 2 shows the intent-to-treat ANCOVA impacts for the six aggregate domains from fully adjusted regressions on the overall sample (pooled waves), as well as at midline and at endline only. The right-most column shows the interaction term between treatment and endline in a pooled model with an interaction term, representing the differential impact for the endline (wave 3), over the midline (wave 2) impact. The pooled model shows impacts on all domains except time use and HIV risk—however for the latter there is an impact observed in just the midline sample. For all other domains, significant impacts accrue in one or more waves. For example, for education, as well as for emotional wellbeing and mental health, there are impacts across both midline and endline (approximately 0.288 – 0.320 for the former and 0.171 – 0.346 SDs over the control mean for the latter). For physical health, impacts accrue primarily at endline (0.206 SDs) and for sexual and reproductive health, impacts accrue primarily at midline (0.095 SDs). However, only in two cases (physical health and emotional wellbeing and mental health) are these differences by wave statistically different from each other. In addition, in the case of physical health, a sensitivity analysis shows that the larger impacts at endline are due to domain construction—as an additional indicator (slept under a mosquito net) was collected at endline (results not shown). Table A13 shows the average impacts for an unadjusted model, as well as full results for the model with covariates. The unadjusted estimates are very similar to the main model, however, the estimates are slightly larger in magnitude and the HIV aggregate becomes significant at the p < 0.10 level, indicating the covariate adjusted model is a conservative lower bound of impacts.

Table 3 replicates these ANCOVA estimates (on the overall sample, using pooled waves) by sex of youth. Although larger absolute effects are found for female youth with respect to physical health, emotional wellbeing and mental and educational domains, the difference is statistically significant only for physical health. In this domain, treatment females have physical health scores that are 0.173 SD larger than males.

4.2. Impacts on individual indicators

Fig. 1 graphically presents the coefficients and 90 % (95 %) confidence intervals for the overall domains, as well as the individual indicators in each domain by sex (aggregate across waves, however full results for individual indicators by wave are reported in Table A14 and by sex in Table A15). For the first domain, physical health, the two main indicators that drive significance of the aggregate index are sleeping under a mosquito net last night (for both male and female youth) and freedom from illness and injury in the last two weeks (female youth). The likelihood of sleeping under a mosquito net increases by about 0.30 SDs compared to youth in the control group (or approximately a 28 % increase in the indicator). For the second domain, mental health and emotional wellbeing, results show that significant impacts on the aggregate index are driven by impacts on youths' freedom from depressive symptoms (0.299 SDs in terms of z-scores, equating to a 9 %

[10] Another way to explicitly address selection is to run analysis using survey weights accounting for the probability of youth being interviewed based on the number of eligible youth in households (following Angeles et al., 2019). We conducted this analysis and found identical results in terms of overall conclusions and similar magnitude of coefficients, thus we conclude that the findings are robust to weighting.

Table 2
ANCOVA impacts on aggregate outcome z-scores with evolution over survey wave.

	Overall effect	N	Effect on midline sample	N	Effect on endline sample	N	Interaction: endline*program impact
Physical health scale (z-score)	0.098	2714	−0.012	1334	0.206	1380	0.227
	(0.057) †		(0.067)		(0.080)*		(0.100)*
Wild bootstrap p-value	*0.143*		*0.882*		*0.029*		*0.028*
Emotional wellbeing & mental health scale (z-score)	0.260	2698	0.171	1325	0.346	1373	0.180
	(0.081)**		(0.090) †		(0.100)**		(0.096) †
Wild bootstrap p-value	*0.001*		*0.099*		*0.003*		*0.078*
Education scale (z-score)	0.301	2715	0.288	1335	0.320	1380	0.009
	(0.065)***		(0.084)**		(0.056)***		(0.071)
Wild bootstrap p-value	*0.000*		*0.004*		*0.000*		*0.885*
Time use scale (z-score)	−0.108	1380	–	–	−0.108	1380	–
	(0.081)		–		(0.081)		–
Wild bootstrap p-value	*0.250*		–		*0.250*		–
Sexual & reproductive health scale (z-score)	0.089	2713	0.095	1333	0.086	1380	−0.000
	(0.047) †		(0.049) †		(0.061)		(0.070)
Wild bootstrap p-value	*0.103*		*0.095*		*0.222*		*0.999*
HIV risk scale (z-score)	0.068	2699	0.124	1326	0.016	1373	−0.131
	(0.044)		(0.041)**		(0.066)		(0.088)
Wild bootstrap p-value	*0.190*		*0.009*		*0.835*		*0.164*

Notes: Standard errors in parenthesis clustered at the district village cluster level. Z-score aggregate outcomes are compiled using standardized indicators, computed in reference to control means. Time use outcomes were not collected at midline. All regressions control for baseline values of the outcome z-score indicator, geographic fixed effects (traditional authorities) as well as a standard set of demographic controls: sex (female) and age of youth (in year fixed effects), sex (female) and age of household head (in continuous years), an indicator for whether the household head ever attended school, as well as household size. † $p < 0.1$ * $p < 0.05$; ** $p < 0.01$; ***$p < 0.001$.

Table 3
ANCOVA impacts on aggregate outcome z-scores by gender of youth.

	Overall effect	N	Effect on male sample	N	Effect on female sample	N	Interaction: female*program impact
Physical health scale (z-score)	0.098	2714	0.029	1445	0.179	1269	0.173
	(0.057) †		(0.070)		(0.068)*		(0.082)*
Wild bootstrap p-value	*0.143*		*0.723*		*0.008*		*0.032*
Emotional wellbeing & mental health scale (z-score)	0.260	2698	0.218	1438	0.307	1260	0.105
	(0.081)**		(0.078)**		(0.119)*		(0.152)
Wild bootstrap p-value	*0.001*		*0.030*		*0.016*		*0.525*
Education scale (z-score)	0.301	2715	0.303	1446	0.304	1269	0.011
	(0.065)***		(0.076)***		(0.096)**		(0.114)
Wild bootstrap p-value	*0.000*		*0.001*		*0.010*		*0.928*
Time use scale (z-score)	−0.108	1380	−0.078	736	−0.121	644	0.002
	(0.081)		(0.086)		(0.155)		(0.203)
Wild bootstrap p-value	*0.250*		*0.423*		*0.487*		*0.996*
Sexual & reproductive health scale (z-score)	0.089	2713	0.070	1444	0.110	1269	0.043
	(0.047) †		(0.058)		(0.069)		(0.111)
Wild bootstrap p-value	*0.103*		*0.289*		*0.151*		*0.706*
HIV risk scale (z-score)	0.068	2699	0.070	1436	0.061	1263	−0.006
	(0.044)		(0.047)		(0.061)		(0.080)
Wild bootstrap p-value	*0.190*		*0.196*		*0.387*		*0.937*

Notes: Standard errors in parenthesis clustered at the district village cluster level. Z-score aggregate outcomes are compiled using standardized indicators, computed in reference to control means. Time use outcomes were not collected at midline. All regressions control for baseline values of the outcome z-score indicator, geographic fixed effects (traditional authorities), as well as a standard set of demographic controls: sex (female, only in fully interacted models) and age of youth (in year fixed effects), sex (female) and age of household head (in continuous years), an indicator for whether the household head ever attended school, as well as household size. † $p < 0.1$ * $p < 0.05$; ** $p < 0.01$; ***$p < 0.001$.

increase in the CES-D scale) and the quality and quantity of social support (0.171 SDs increase in z-scores). For the former, impacts are larger in magnitude for the female sample, while the reverse is true for the latter—however in both cases the difference is not statistically different. There are no impacts on aspirations across either sex. For the third domain, education, overall impacts are driven by school enrollment (0.268 SDs in terms of z-scores; equating to a 26 % increase in the underlying indicator) and education expenditure (0.327 SDs in terms of z-scores; equating to a 79 % increase in the underlying indicator). Despite the increased enrollment, we find no significant impacts on school attendance in the last week (among those currently attending), or on the highest grade completed. For the fourth domain, although we find no overall impacts on the aggregate time use indicator, we find that the intervention significantly increases the likelihood of youth engaging in hazardous labor at endline (an indicator which included hazards of: a) carrying heavy loads, b) working with dangerous tools, c) exposure to

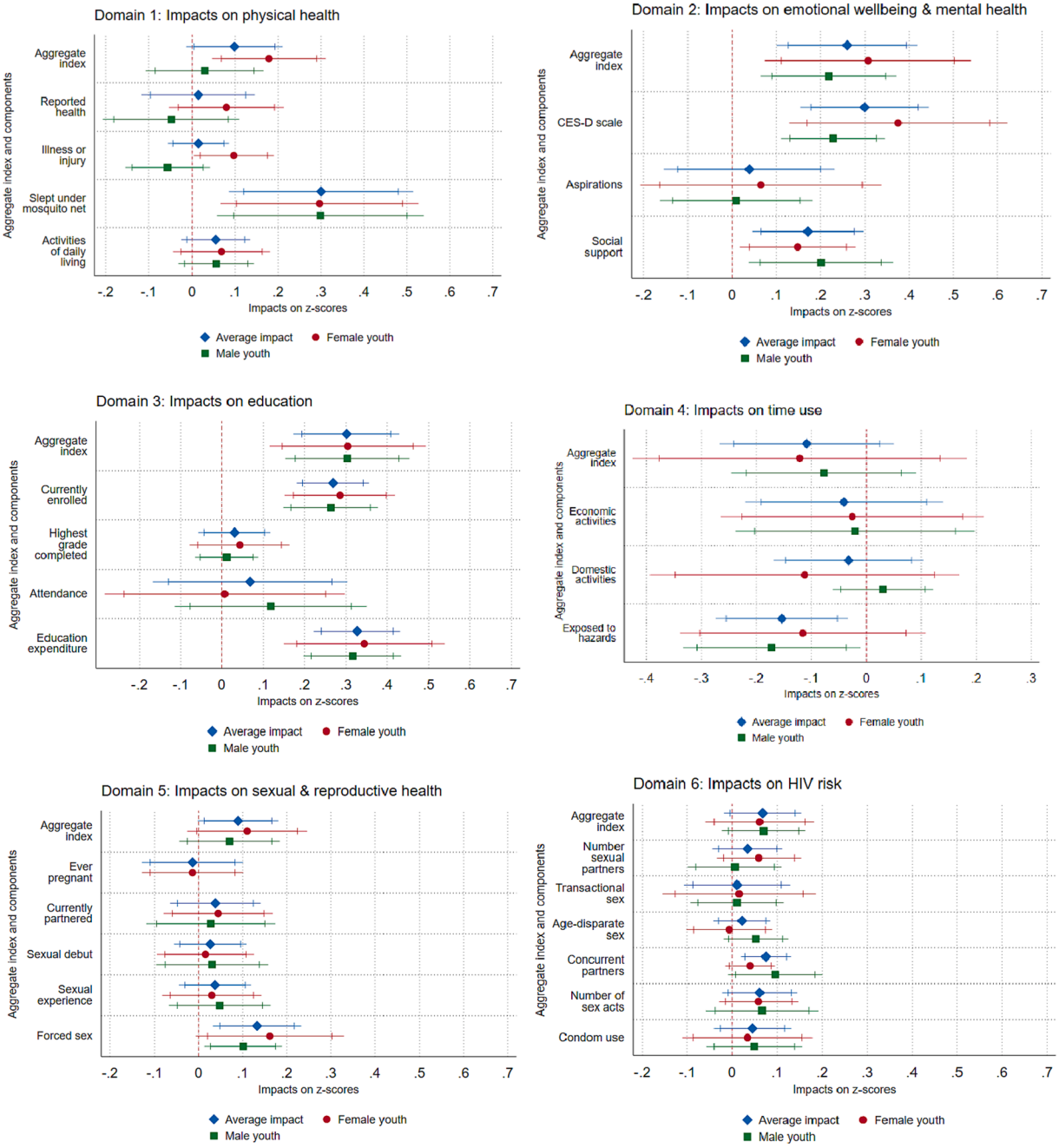

Fig. 1. ANCOVA impacts on domains and indicator z-scores by gender (aggregate across waves). *Notes:* 90% and 95% confidence intervals are shown in bars (standard errors are not corrected for number of clusters or multiple hypothesis testing). All indicators are revised so higher z-scores equal favorable outcomes. Standard errors are clustered at the district village cluster level. Z-score aggregate outcomes are compiled using standardized indicators, computed in reference to control means. All regressions control for baseline values of the outcome z-score indicator, geographic fixed effects (traditional authorities), as well as a standard set of demographic controls: sex (female) and age of youth (in year fixed effects), sex (female) and age of household head (in continuous years), an indicator for whether the household head ever attended school, as well as household size. Full results are given in Table A15.

dust, fumes or gas or d) exposure to extreme heat, cold or humidity). There are few impacts on indicators in the fifth and sixth domains, with a few exceptions. For sexual and reproductive health, the program has beneficial impacts on both male and female youth reporting of being forced, tricked or pressured into sex (overall reducing reporting of never experiencing acts by 0.133 SDs in terms of z-scores, equivalent to a 9 % increase in the underlying indicator; 0.096 SDs for male youth and 0.161 SDs for female youth). In addition, the program increases the incidence of no partner concurrency (0.075 SDs), a beneficial change for HIV risk, driven by male youth.

4.3. Robustness checks

We conduct a number of robustness checks to show the sensitivity of our results to alternative specifications. First, we replicate the main estimates shown in Table 2 using DD models and present these in

Table A16.[11] Overall findings are similar to ANCOVA estimates although some differences are noted. For example, while ANCOVA estimates show significant impacts on both physical health and emotional and mental well-being domains, the DD estimates are similar in magnitude, but not significant. These results are consistent with the improved power due to ANCOVA modeling. Nonetheless, estimates in both models support the evaluation of impacts overtime for physical health and mental health (with significant interaction effects in both cases of similar magnitude). Both models also produce similar findings for education and time use domains. Finally, for sexual and reproductive health and HIV risk scales, DD estimates produce higher significance levels and magnitude of coefficients. In DD models, there are average impacts for sexual and reproductive health of 0.222 SD and HIV risk of 0.188 SDs (both significant at the 1 % level)—which appear to be stable over time. In this sense, ANCOVA estimates appear to be a lower bound for both sexual and reproductive health and HIV risk.

Second, we acknowledge that holistic measures of safe transition encompass many different outcomes—thus there may be concerns that the analysis suffers from false positive results due to many hypothesis tests. As explained in previous sections, we partially address this issue by presenting analysis of aggregates containing all possible relevant indicators collected as part of the survey effort. However, to address concerns for the individual outcome indicators, we present q-values that correct for false discovery rates for average impacts across waves in Table A17 (Menyhart et al., 2021). For this analysis, we use the more conservative p-values adjusting for the small number of clusters in the study. Results show that, although significance levels are reduced, considering a cut-off of 10 % significance, only one (no hazardous labor) becomes insignificant (p-value = 0.046, q-value = 0.114). This indicates that adverse effects increasing the incidence of hazardous labor may not hold when considering multiple hypothesis testing. For all other outcomes showing individual significance (seven in total), beneficial significant impacts of the cash transfer appear to hold.

Third, because more than one youth may reside in any given household, we replicate overall average effects (shown in Table 2) controlling for household fixed effects—as youth living in the same household may be subject to underlying similarities that may affect their outcomes. As shown in Fig. A2, there are an average of 1.3 youth per household in the analysis sample, with similar rates for treatment and control households. Panel A of Table A18 shows that outcomes controlling for household fixed effects are very similar to those in main models, thus results are robust to accounting for underlying fixed household characteristics.

Finally, despite little evidence of differential attrition, the overall rates of attrition for youth in our study are meaningful (20 percent). These rates are in line with a review of RCTs from LMICs from economic journals, showing attrition rates of individual children (18 years and younger) was 11.6 percent (range: 0.6 to 60 percent), while those for adults (above 18 years) was 19.4 percent (range: 1.3 to 75 percent) for unreported study periods (Millán & Macours, 2017). Nonetheless, we use the methodology proposed by (Lee, 2009) to estimate treatment-effect bounds using the command developed for Stata (Tauchmann, 2014). We report these using the combined midline and endline single difference follow-ups in Panel B of Table A18. We control for district TA only in this analysis as the 'leebound' command faces convergence issues with many covariates and does not allow for continuous covariates (preventing us from including baseline values of outcome indices). While lower and upper bounds mostly mirror the significance levels for our main coefficients, several of the lower bounds cross zero (those for time use and HIV risk). For time use specifically, the lower bound crosses zero and is statistically significant (at the $p < 0.05$ level). This is expected given the negative average coefficient in our main models and the negative significant sign of several individual time use indicators as noted in the previous section.

5. Discussion and conclusion

As government-run social safety nets are increasingly implemented in Africa and globally, evidence on the impact on safe youth transitions of programs can help guide investment and program design decisions. The current paper shows that youth living in households experiencing ultra-poverty improved outcomes after one and two years of household-receipt of transfers. These improvements are not only in domains thought to be directly connected to influx of income (e.g., education, physical health) but also across additional domains of emotional well-being and mental health, as well as sexual and reproductive health. We find particularly strong results on indicators of sleeping under a mosquito net, freedom from illness and injury, school enrollment and expenditures, social support, freedom from depressive symptoms and from being forced, tricked or pressured into sex. While our results cannot be strictly compared to other evaluations—the magnitude of average impacts, ranging from 0.089 to 0.300 SDs, is in line with other studies on short-term impacts. For example, the analysis of a conditional cash transfer to secondary school-aged girls in South Africa found impacts on domains ranging from 2 percentage points (pp) (relationships) to 9 pp (economic agency) after three years, aggregated using the multi-dimensional overlapping deprivation analysis approach (Kilburn et al., 2020). An evaluation of the "Girl Empower" intervention in Liberia, combining transformative mentoring and cash transfers (targeting 13–14-year-olds at baseline) found impacts of the combined intervention on aggregate indices of sexual and reproductive health (0.372 SDs), gender attitudes (0.228 SDs) and life skills (0.289 SDs) after 24 months—however no impacts on aggregates for sexual violence, psychosocial or protective factors (Özler et al., 2020). Both of these evaluations focused exclusively on female youth; however, we show that there are equally promising opportunities for cash transfers to improve outcomes for male youth. These results confirm favorable impacts from broader reviews and confirm findings from correlative studies that argue in favor of cash transfers as a promising 'accelerator' for the SDGs for at risk African youth (Cirillo et al., 2021; Cluver et al., 2019; Peterman et al., 2017; Stoner et al., 2021; Zimmerman et al., 2021).

While we do not conduct a formal analysis of mechanisms underlying these impacts—it is likely that many are driven by an economic or income effect. This hypothesis is supported by analysis showing the Malawi SCTP increased consumption, food security, asset ownership and resilience at the household level (Handa et al., 2022; Otchere & Handa, 2022). Increases in available resources could link directly to youth outcomes (e.g., via purchase and use of mosquito nets, or school expenditures), as well as foster improvements in health or wellbeing for household members in general. As shown by qualitative work within the study sample, families are cohesive economic and social support units, thus it is likely that youth outcomes are also affected positively or negatively by other household and community-member's behaviors (Barrington et al., 2017; Rock et al., 2016). It is also possible that sequencing occurs in impacts—with some impacts occurring first, which then affect subsequent outcomes—acting as mechanisms of change. For example, the aforementioned analysis on mental health outcomes of female youth from this evaluation showed that increases in education, social support and reductions in caregiver stress and youth engagement in casual labor were salient mediators of program impact (Angeles et al., 2019). It is likely that mechanisms of impact are complex and interrelated, and that qualitative work is needed to holistically unpack

[11] The DD specification replicates equation (1) for the six domains: $Y_{ihjt} = \beta_0 + \beta_1 Treat_j * Endline_{it} + \beta_2 Treat_j + Endline_{it} + \beta_3 X_{ihj0} + \gamma_{TA} + \varepsilon_{ihj1}$ Where $Endline_{it}$ is a dummy variable which takes on the value of 1 if the individual is observed at endline and 0 otherwise. Similar to equation (1), the vector β_1 captures the program impact on the various outcomes of interest. $X_{hij0}, \gamma and \varepsilon_{ihj1}$ remain the same as explained for equation (1). For heterogeneous impacts using the DD specification, we implement a triple interaction model, as well as separate regressions for male and female (midline and endline) samples.

these dynamics, alongside other outcomes which may be important to youth transitions, including pride, dignity and transitions to productivity.

Our results support Cirillo and colleagues (2021) who argue that social protection can contribute to a 'pipeline of investments' and help support at-risk adolescents—ultimately facilitating a triple dividend as youth transition to adulthood and parenthood, nurturing the next generation. Depending on the setting and target population, Cirillo and colleagues (2021) suggest these risks include not only poverty and economic insecurity, but also disability, chronic illness, and social discrimination based on gender, race, or ethnicity. In addition, authors argue social protection has the potential to play a transformative role, facilitating increased agency, autonomy and empowerment. They distinguish between two main types of programs in relation to their potential impacts on adolescents. The first are programs that target households experiencing poverty more generally but have program design or implementation components that are sensitive to adolescent needs (e.g., target households that are likely to have adolescents, give larger transfers based on presence of an adolescent or have conditionalities related to secondary schooling). The second are programs that target and cater to adolescents specifically. In a mapping exercise of government programs, they show that most programs fall into the former category—however there is a large variation in design components that may make programs adolescent-sensitive. While the Malawi SCTP can be classified in the first category, it also has elements of an 'adolescent-sensitive' approach. These include the stated objective of improving school enrollment and providing larger transfers to households with secondary-school age children. Despite this, we find one adverse impact, on the likelihood of youth engaging in hazardous labor. This result has also been documented using data from the household survey of the same evaluation among school-aged children and explained in the context of increased household-level economic activities that pulled youth into productive labor (de Hoop et al., 2020). Therefore, additional program design or operational elements may be needed to ensure that adverse impacts for youth are minimized and trade-offs for overall wellbeing are carefully considered.

There are several limitations of these results that are worth mentioning. First, the total number of clusters in the evaluation is considered low (n = 29). We provide corrections to standard errors to account for this limitation; however, we acknowledge that some analysis may have been underpowered to detect impacts. In particular, our heterogenous impacts by gender and survey wave should be taken with caveats, as the study may be underpowered to detect differences across these categories. Second, while the study included three years (waves) of data, several key outcomes, including pregnancy and marriage, are outcomes expected to occur over a longer time span. Thus, it is possible that some specific outcomes need a longer time horizon to show full effects. Finally, the evaluation was set up to follow households over time—rather than youth themselves. Therefore, attrition was higher than would be expected in a longitudinal evaluation that explicitly followed youth. Some of these youth may have left households for reasons that may have implications for the outcomes we focus on in the current study. For example, youth may leave households to marry and set up their own households, or to seek employment or income earning opportunities in urban centers. Finally, ideally, we would be able to assess how youth of different characteristics responded differentially to the intervention (e.g. youth with disabilities, youth who are orphans). However, due to the sample size requirements for analyzing heterogenous effects, we are not confidently able to conduct this analysis.

Although promising, our results suggest that more can be done to intentionally design for gender-sensitive impacts on youth safe transitions. First, although the current transfer value calculation includes higher values for secondary school-age youth—this formula could be revisited to accommodate the transition out of secondary school, into employment or vocational training. Second, given that girls are more likely to be out of school, additional bonuses could be given to households with female secondary-school age members. These accommodations could be accompanied by a messaging campaign that 'labels' these bonuses for youth investment and which emphasize the importance of keeping girls in school. Labelled approaches have the benefit of encouraging behaviors, without strict conditions which are likely to leave out or exclude the most vulnerable households and youth (Benhassine et al., 2015). Third, the program should consider if a rolling application or eligibility process could be put in place to facilitate the enrollment of youth as primary beneficiaries once they become adults and heads of their own households. Finally, the program could actively seek out ways to offer adolescents who are out of school the ability to take part in other economic programs (e.g., training, mentorship or apprentice activities) or link them to services (e.g., youth-friendly sexual and reproductive health services)—designed specifically to tackle some of the multi-dimensional ongoing challenges to safe transitions. These recommendations build on the ongoing range of innovative and promising programmatic models being tested to accelerate outcomes for youth on the African continent and beyond (Cirillo et al., 2021). Finally, additional mixed-methods research is needed on social safety nets and youth, not just on design considerations, but also focusing on gender differences and longer-term impacts. This includes the need for future research on cost-effectiveness focusing on holistic benefits for youth and intergenerational outcomes to guide policy makers on the optimal required investments.

CRediT authorship contribution statement

Monica Lambon-Quayefio: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Amber Peterman:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Sudhanshu Handa:** Conceptualization, Methodology, Writing – review & editing, Investigation, Supervision, Project administration. **Adria Molotsky:** Formal analysis, Investigation, Data curation, Writing – review & editing. **Frank Otchere:** Investigation, Data curation, Writing – review & editing. **Peter Mvula:** Conceptualization, Investigation, Supervision, Funding acquisition, Writing – review & editing, Project administration. **Maxton Tsoka:** Conceptualization, Investigation, Supervision, Funding acquisition, Writing – review & editing, Project administration. **Jacobus de Hoop:** Conceptualization, Investigation, Data curation, Writing – review & editing. **Gustavo Angeles:** Conceptualization, Methodology, Investigation, Funding acquisition, Supervision, Writing – review & editing. **Kelly Kilburn:** Investigation, Data curation, Writing – review & editing. **Annamaria Milazzo:** Methodology, Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

We thank the Malawi Social Cash Transfer Evaluation Team for their contribution to this study. In addition, our appreciation goes to the Government of Malawi for their supportive engagement with the evaluation team, and for their time and intellectual contributions, specifically Dr. Mary Shawa, Dr. Esmie Kainja, Mr. Laurent Kansinjiro, Mr. Charles Chabuka and Mr. Gideon Kachingwe of the Ministry of Gender, Children, Disabilities and Social Welfare, Mr. Harry Mwamlima of the Ministry of Finance, Economic Planning and Development, as well as the

District Commissioner's Offices of Salima and Mangochi. We would also like to acknowledge Chantal Elmont of Ayala Consulting, who continually supported the evaluation team by providing vital information on the implementation of the SCTP. Our thanks go to the field research team at CSR for their exceptional work and all the youth and households who took part in the data collection. We thank the European Union, the German Government through KfW, Irish Aid, FAO, the International Initiative for Impact Evaluation (3ie), the UK Foreign, Commonwealth and Development Office and UNICEF Malawi for their financial contributions and stakeholder support for the study. We also thank the William and Flora Hewlett Foundation for support of the Transfer Project early career research fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.worlddev.2023.106483.

References

Angeles, G., de Hoop, J., Handa, S., Kilburn, K., Milazzo, A., & Peterman, A. (2019). Government of Malawi's unconditional cash transfer improves youth mental health. *Social Science & Medicine, 225*, 108–119. https://doi.org/10.1016/j.socscimed.2019.01.037

Baird, S., Garfein, R. S., McIntosh, C. T., & Özler, B. (2012). Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: A cluster randomised trial. *The Lancet, 379*(9823), 1320–1329. https://doi.org/10.1016/S0140-6736(11)61709-1

Baird, S., McIntosh, C., & Özler, B. (2011). Cash or Condition? Evidence from a Cash Transfer Experiment. *The Quarterly Journal of Economics, 126*(4), 1709–1753. https://doi.org/10.1093/qje/qjr032

Baird, S., McIntosh, C., & Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics, 140*, 169–185. https://doi.org/10.1016/j.jdeveco.2019.04.004

Barrington, C., Villa-Torres, L., Abdoulayi, S., Tsoka, M. G., & Mvula, P. M. (2017). Using Photo-Elicitation Methods to Understand Resilience Among Ultra-Poor Youth and Their Caregivers in Malawi. *Health Education & Behavior, 44*(5), 758–768. https://doi.org/10.1177/1090198117728756

Beegle, K., Coudouel, A., & Monsalve, E. (2018). Realizing the Full Potential of Social Safety Nets in Africa. *The World Bank*. https://doi.org/10.1596/978-1-4648-1164-7

Benhassine, N., Devoto, F., Duflo, E., Dupas, P., & Pouliquen, V. (2015). Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education. *American Economic Journal: Economic Policy, 7*(3), 86–125. https://doi.org/10.1257/pol.20130225

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics, 90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Carolina Population Center. (2014). *Malawi Social Cash Transfer Program: Baseline Evaluation Report*. https://transfer.cpc.unc.edu/wp-content/uploads/2021/04/Malawi-SCTP-Baseline-Report.pdf.

Carolina Population Center. (2016). *Malawi Social Cash Transfer Programme: Endline Impact Evaluation Report*. https://transfer.cpc.unc.edu/wp-content/uploads/2021/04/Malawi-SCTP-Endline-Report_Final.pdf.

Cirillo, C., Palermo, T., & Viola, F. (2021). *Non-contributory Social Protection and Adolescents in Lower- and Middle-Income Countries: A review of government programming and impacts* (Office of Research–Innocenti Working Paper 2021–07). https://www.unicef-irc.org/publications/1245-non-contributory-social-protection-and-adolescents-in-lower-and-middle-income-countries-a-review-of-government-programming.html.

Cluver, L. D., Orkin, F. M., Campeau, L., Toska, E., Webb, D., Carlqvist, A., & Sherr, L. (2019). Improving lives by accelerating progress towards the UN Sustainable Development Goals for adolescents living with HIV: A prospective cohort study. *The Lancet Child & Adolescent Health, 3*(4), 245–254. https://doi.org/10.1016/S2352-4642(19)30033-1

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature, 554*(7693), Article 7693. doi: 10.1038/nature25770.

Dake, F., Natali, L., Angeles, G., de Hoop, J., Handa, S., Peterman, A., & on behalf of the Malawi Cash Transfer Evaluation Team and the Zambia Cash Transfer Evaluation Team. (2018). Cash Transfers, Early Marriage, and Fertility in Malawi and Zambia. *Studies in Family Planning, 49*(4), 295–317. doi: 10.1111/sifp.12073.

de Hoop, J., Groppo, V., Handa, S., & on behalf of the Malawi Social Cash Transfer Program and the Zambia Multiple Category Targeted Program study teams. (2020). Cash Transfers, Microentrepreneurial Activity, and Child Work: Evidence from Malawi and Zambia. *The World Bank Economic Review, 34*(3), 670–697. doi: 10.1093/wber/lhz004.

de Walque, D., Fernald, L., Gertler, P., & Hidrobo, M. (2017). Cash Transfers and Child and Adolescent Development. In D. A. P. Bundy, N. de Silva, S. Horton, D. T. Jamison, & G. C. Patton (Eds.), *Child and Adolescent Health and Development* (3rd ed.). The International Bank for Reconstruction and Development / The World Bank. http://www.ncbi.nlm.nih.gov/books/NBK525264/.

GAGE Consortium. (2017). *Gender and Adolescence: Why understanding adolescent capabilities, change strategies and context matters*. https://www.gage.odi.org/wp-content/uploads/2019/01/Gender-and-Adolescence-CF-FINAL-1.pdf.

Handa, S., Kilburn, K., Barrington, C., Angeles, G., Mvula, P., & Tsoka, M. (2013). *Impact Evaluation of the Malawi SCT: Inception Report*. Carolina Population Center: University of North Carolina.

Handa, S., Otchere, F., Sirma, P., & Team, the E. S. (2022). More evidence on the impact of government social protection in sub-Saharan Africa: Ghana, Malawi, and Zimbabwe. *Development Policy Review, 40*(3), e12576. doi: 10.1111/dpr.12576.

Handa, S., Peterman, A., Huang, C., Halpern, C., Pettifor, A., & Thirumurthy, H. (2015). Impact of the Kenya Cash Transfer for Orphans and Vulnerable Children on early pregnancy and marriage of adolescent girls. *Social Science & Medicine, 141*, 36–45. https://doi.org/10.1016/j.socscimed.2015.07.024

Heinrich, C. J., Hoddinott, J., & Samson, M. (2017). Reducing Adolescent Risky Behaviors in a High-Risk Context: The Effects of Unconditional Cash Transfers in South Africa. *Economic Development and Cultural Change, 65*(4), 619–652. https://doi.org/10.1086/691552

Kilburn, K., Ferrone, L., Pettifor, A., Wagner, R., Gómez-Olivé, F. X., & Kahn, K. (2020). The Impact of a Conditional Cash Transfer on Multidimensional Deprivation of Young Women: Evidence from South Africa's HTPN 068. *Social Indicators Research, 151*(3), 865–895. https://doi.org/10.1007/s11205-020-02367-y

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica, 75*(1), 83–119. https://doi.org/10.1111/j.1468-0262.2007.00733.x

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies, 76*(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics, 99*(2), 210–221. https://doi.org/10.1016/j.jdeveco.2012.01.002

Menyhart, O., Weltz, B., & Győrffy, B. (2021). MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction. *PLOS ONE, 16*(6), e0245824.

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2019). Long-Term Impacts of Conditional Cash Transfers: Review of the Evidence. *The World Bank Research Observer, 34*(1), 119–159. https://doi.org/10.1093/wbro/lky005

Millán, T. M., & Macours, K. (2017). *Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias*. https://www.iza.org/publications/dp/10711/attrition-in-randomized-control-trials-using-tracking-information-to-correct-bias.

National Statistics Office. (2014). *Integrated Household Panel Survey 2010-2013: Household Socio-Economic Characteristics Report*.

Otchere, F., & Handa, S. (2022). Building Resilience through Social Protection: Evidence from Malawi. *The Journal of Development Studies, 58*(10), 1958–1980. https://doi.org/10.1080/00220388.2022.2075733

Özler, B., Hallman, K., Guimond, M.-F., Kelvin, E. A., Rogers, M., & Karnley, E. (2020). Girl Empower – A gender transformative mentoring and cash transfer intervention to promote adolescent wellbeing: Impact findings from a cluster-randomized controlled trial in Liberia. *SSM - Population Health, 10*, Article 100527. https://doi.org/10.1016/j.ssmph.2019.100527

Parker, S. W., & Vogl, T. (2018). *Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico* (Working Paper 24303). National Bureau of Economic Research. doi: 10.3386/w24303.

Peterman, A., Neijhoft, A. (Naomi), Cook, S., & Palermo, T. M. (2017). Understanding the linkages between social safety nets and childhood violence: A review of the evidence from low- and middle-income countries. *Health Policy and Planning, 32*(7), 1049–1071. doi: 10.1093/heapol/czx033.

Pettifor, A., MacPhail, C., Hughes, J. P., Selin, A., Wang, J., Gómez-Olivé, F. X., … Kahn, K. (2016). The effect of a conditional cash transfer on HIV incidence in young women in rural South Africa (HPTN 068): A phase 3, randomised controlled trial. *The Lancet Global Health, 4*(12), e978–e988. https://doi.org/10.1016/S2214-109X(16)30253-4

Rock, A., Barrington, C., Abdoulayi, S., Tsoka, M., Mvula, P., & Handa, S. (2016). Social networks, social participation, and health among youth living in extreme poverty in rural Malawi. *Social Science & Medicine, 170*, 55–62. https://doi.org/10.1016/j.socscimed.2016.10.005

Roodman, D., Nielsen, M.Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal, 19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Stoner, M. C. D., Kilburn, K., Godfrey-Faussett, P., Ghys, P., & Pettifor, A. E. (2021). Cash transfers for HIV prevention: A systematic review. *PLOS Medicine, 18*(11), e1003866.

Tauchmann, H. (2014). Lee (2009) Treatment-Effect Bounds for Nonrandom Sample Selection. *The Stata Journal, 14*(4), 884–894. https://doi.org/10.1177/1536867X1401400411

United Nations. (2020). *World Population Prospects 2019 - Volume II: Demographic Profiles*. United Nations. doi: 10.18356/7707d011-en.

Zimmerman, A., Garman, E., Avendano-Pabon, M., Araya, R., Evans-Lacko, S., McDaid, D., … Lund, C. (2021). The impact of cash transfers on mental health in children and young people in low-income and middle-income countries: A systematic review and meta-analysis. *BMJ Global Health, 6*(4), e004661.