Public Disclosure Authorized

Public Disclosure Authorized

WPS5644

Policy Research Working Paper 5644

Income Shocks and Adolescent Mental Health

Sarah Baird
Jacobus de Hoop
Berk Özler

The World Bank
Development Research Group
Poverty and Inequality Team
April 2011

Policy Research Working Paper 5644

Abstract

In this paper, the authors investigate the effect of positive income shocks on the mental health of adolescent girls using experimental evidence from a cash transfer program in Malawi. They find that the provision of monthly cash transfers had a strong beneficial impact on the mental health of school-age girls during the two-year intervention. Among baseline schoolgirls who were offered unconditional cash transfers, the likelihood of suffering from psychological distress was 38 percent lower than the control group, while the same figure was 17 percent if the cash transfers offers were made conditional on regular school attendance. The authors find no impact on the mental health of girls who had already dropped out of school at baseline. The beneficial effects of cash transfers were limited to the intervention period and dissipated quickly after the program ended.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at bozler@worldbank.org.

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

Produced by the Research Support Team

INCOME SHOCKS AND ADOLESCENT MENTAL HEALTH[1]

Sarah Baird, George Washington University
Jacobus de Hoop, Tinbergen Institute and VU University Amsterdam
Berk Özler, World Bank

Keywords: adolescent girls; cash transfers; mental health; randomized intervention.

JEL Codes: C93; I10; I21; I38.

[1] We are grateful to participants at the National Bureau of Economic Research (NBER) Africa Project Research Conference and participants at a seminar in the University of Oxford for valuable comments. We gratefully acknowledge funding from the Global Development Network, the Bill and Melinda Gates Foundation, NBER, World Bank Research Support Budget Grant, as well as several trust funds at the World Bank: Knowledge for Change Trust Fund (TF090932), World Development Report 2007 Small Grants Fund (TF055926), Spanish Impact Evaluation Fund (TF092384), and Gender Action Plan Trust Fund (TF092029). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development or the World Bank. Please send correspondence to: sbaird@gwu.edu, jdhoop@pse.ens.fr, or bozler@worldbank.org.

1. INTRODUCTION

Mental health problems make the leading contribution to the global disease and injury burden among adolescent women (World Health Organization 2004).[2] Mental health problems developed during adolescence are not only of intrinsic immediate importance to young women, but they can also have severe long-run health consequences. They often persist into adulthood, increase the risk of contracting communicable and non-communicable diseases, and are strongly associated with reproductive and sexual health (Patel et. al., 2007; Prince et. al., 2007). In addition, mental health problems are linked to adverse developmental outcomes, such as lower educational achievement (see, e.g., Currie and Stabile, 2006; Eisenberg and Golberstein, 2009; Fletcher, 2008; Fletcher, 2010; Fletcher and Wolfe, 2008; Kessler et. al. 1995; Stein and Kean, 2000). These adverse health and developmental effects can in turn have a long-term impact on socioeconomic outcomes and thus result in a vicious circle of poverty and poor mental health (Patel and Kleinman 2003).

Das et. al. (2007; 2008) suggest that unexpected life events (either positive or negative shocks) can have a strong influence on mental health. Friedman and Thomas (2008) and Stillman, McKenzie, and Gibson (2009) are prominent examples of studies confirming the impact of shocks on mental health. Friedman and Thomas (2008) show that the Indonesian financial crisis of the 1990s had a detrimental effect on the psychological wellbeing of the Indonesian people. Stillman, McKenzie, and Gibson (2009), on the other hand, show that individuals who were selected by lottery to migrate from Tonga to New Zealand (which has higher living standards) exhibited significantly improved mental health outcomes.

[2] In a review of 11 epidemiological community studies from developed and developing countries, Patel et. al. (2007) show that the prevalence rate of mental disorders among adolescents ranges from 8% in the Netherlands to 27% in Australia. Women are more likely than men to suffer from poor mental health and depressive disorders (Das et. al. 2008; Patel et. al. 2007). This gender differential is likely to be the result of a combination of genetic and environmental risk factors (Patel and Kleinman 2003). Poor mental health also demands a heavy toll among young women in Sub-Saharan Africa: after HIV/AIDS and abortion, depression makes up the leading contribution to disability adjusted life-years (DALYs) among this demographic group.

As the Indonesian financial crisis and Tonga-New Zealand migration entailed large negative and positive income changes respectively, these studies appear to suggest that income shocks may play an important role in mental health outcomes. However, since both of these events also led to many other covariate changes, it is impossible to directly identify the effect of income shocks on mental health from these studies. Gardner and Oswald (2007) provide more direct evidence on the impact of positive income shocks on mental wellbeing by showing that British lottery winners exhibit significant improvements in mental health. However, as Gardner and Oswald (2007) note, their results suffer from a potential endogeneity concern, as it cannot be ruled out that there are structural differences between lottery winners and the individuals in their control group.[3]

This paper uses a randomized cash transfer experiment to identify the causal effect of positive income shocks on the mental health of school-age girls in Malawi. Considering the importance of mental health during adolescence in educational achievement and risky behaviors, school-age beneficiaries are a target group of particular interest.[4] The unique design of the intervention allows us to compare the impact of *unconditional* cash transfers (i.e. pure income shocks) with the impact of cash transfers made *conditional* on attending school. Multiple overlapping randomized treatment layers also allow us to investigate the elasticity of mental health with respect to transfer size, to examine whether the identity of the recipient matters, and to identify spillover

[3] There is much more evidence on the relationship between income *levels* and mental wellbeing. This literature was given impetus by Easterlin (1974). He showed that within countries there is a link between income and happiness. However, across countries average levels of happiness do not exhibit much variation with respect to income. For references to this extensive literature we refer the reader to the introduction of Gardner and Oswald (2007).

[4] There is a fair amount of evidence on the causes of poor mental health among adolescents. Patel et. al. (2007) identify a range of risk factors including poverty and social disadvantage, unstable romantic relationships, violence and abuse, poor physical health, and inadequate education. Das et. al. (2008) are unable to confirm a relationship between poverty and mental health using survey data on mental health from 5 developing countries. Engaging family and education environments (such as parents encouraging children to express their emotions and schools providing a safe learning atmosphere), on the other hand, can reduce the probability that adolescents develop mental disorders (Lewinsohn, Rohde and Seeley 1998; Patel et. al. 2007; Saluja et. al. 2004).

effects separately on school-age siblings and on other school-age girls within treatment communities.[5]

Using the General Health Questionnaire 12 (GHQ-12), a screening instrument widely used in clinical settings, we show that a positive income shock can cause a substantial temporary reduction in psychological distress. While the intervention was ongoing, baseline schoolgirls who were offered unconditional cash transfers were approximately 14 percentage points (pp), or 38%, less likely to be classified as suffering from mental health problems. Those who had been offered transfers *conditional* on regular school attendance also experienced a statistically significant improvement in their mental health, but at roughly 6 pp (17%) the size of this impact is significantly lower than that for the unconditional treatment group. The likelihood that a baseline schoolgirl suffers from mental health problems increased by approximately 3 pp with each additional dollar offered to her parents *conditional* on her school attendance, which is responsible for this muted impact in the conditional treatment arm. The intervention resulted in detrimental spillover effects on the mental health of girls who lived in a treatment area but were not invited to participate in the intervention and in positive spillover effects on untreated female siblings of girls who had been offered a transfer. The impact of the intervention on mental health does not appear to be permanent, as the measured differences between the treatment and control groups disappeared to a large extent after the intervention ended.

[5] There is now a substantial literature examining the impact of cash transfer interventions on physical health outcomes (See Fiszbein and Schady, 2009, for a review). Evidence is accumulating on the impact of cash transfers on physical health outcomes for infants and children under the age of six, such as stunting, anemia, and body-mass (e.g. Fernald, Hou and Gertler 2008; Fernald, Gertler and Neufeld 2009; Gertler 2004; Lagarde, Haines and Palmer 2007; Rivera et. al. 2004). There is also evidence on the impact of cash transfer interventions on behavior, cognition, and stress levels in infants and young children in Latin America (e.g. Fernald and Gunnar 2009; Fernald, Gertler and Neufeld 2009; Macours, Schady and Vakis 2008; Ozer et. al. 2009; Paxson and Schady 2007). Filmer and Schady (2009) present evidence that a cash transfer intervention in Cambodia had a small effect on the mental health of its adolescent beneficiaries. Except for this latter study, however, there is little evidence on the impact of cash transfer interventions on the mental health of adolescents.

The remainder of this paper is as follows. Section 2 discusses the instrument we use to measure mental health. Section 3 describes the study setting, the research design and the intervention. Section 4 describes our estimation strategy and presents the results. Section 5 concludes.

2. MENTAL HEALTH INSTRUMENT

Our main instrument to measure mental health problems is the GHQ-12. The GHQ-12 is a screening instrument widely used in clinical settings to identify psychological distress and common mental disorders. The instrument contains twelve Likert items, each related to mental health.[6] Respondents who report an item as applying to them "more than usual" or "much more than usual" ("less than usual" or "much less than usual" for positively phrased items) score 1 on that item, and 0 otherwise. The summed scores, ranging from 0 to 12, provide an indication of the extent to which a respondent suffers from mental health problems. Summed scores with a value of 3 or higher are typically classified as cases suffering from psychological distress or common mental disorders (caseness), such as anxiety and depression (Weich and Lewis 1998; Weich, Lewis, and Jenkins 2001; Wiggins et. al. 2004). This threshold score of 3, however, is not universal.[7] The correct threshold score is context dependent and can be determined through a validation procedure which compares scores on the GHQ-12 to caseness determined by means of in-depth psychiatric interviews. We did not validate the GHQ-12 in Malawi using such a procedure.

The GHQ-12 was designed for use in clinical settings, but it is regularly used to screen for psychological distress in non-clinical settings. Examples include Weich and Lewis (1998), Weich,

[6] We use a version of the GHQ-12 with five-level Likert items. The Likert items of the standard GHQ-12 contain the following four levels: "much more than usual", "more than usual", "same as usual", and "not at all". The version we use replaces the last level of the standard version with "less than usual", and "much less than usual".
[7] Some studies suggest using a threshold score of 4 (e.g. Jackson, 2007).

Lewis and Jenkins (2001) and Wiggins et. al. (2003), all of which investigate the relationship between socioeconomic indicators and mental health problems in the U.K. There are two important arguments in support of using the GHQ-12 as an outcome measure in the non-clinical setting of this study. First, the GHQ-12 has been shown to be reliable in a wide range of cultural contexts. The GHQ-12 was originally drafted in English by Goldberger (1988), but has been shown to be reliable and have high sensitivity and specificity in different settings. It has been translated to, and validated in, all major languages and many smaller languages such as Malay (Yusoff, Rahim and Yaacob 2009), Persian (Montazeri et. al. 2003)**,** and Yoruba (Gureje and Obikoya 1990). For this study the GHQ-12 was translated into Chichewa, the main language spoken in the study area. Second, while the GHQ-12 was originally developed as an instrument to detect psychological distress in adults, it has also been shown to be a reliable instrument among adolescents (Tait, French, and Hulse 2003; French and Tait 2004; Montazeri et. al. 2003).

Although the GHQ-12 is mostly regarded as providing a general measure of mental health and assessing the likelihood of suffering from psychological distress, it has been suggested that it can also be used as a multidimensional measure of mental health. In an influential paper, Graetz (1990) identified three underlying dimensions of mental health measured by the GHQ-12 by looking at items that behave uniformly in response to outside variables. Graetz (1990) named these dimensions "anxiety and depression" (four items), "social dysfunction" (six items), and "loss of confidence" (two items). The underlying dimensions identified in that paper have since been confirmed in a number of studies (e.g. Gao et. al. 2004; Ye 2009). Appendix 1 lists the GHQ-12 questions, categorized into these three dimensions.

We use the binary indicator of psychological distress, which is equal to unity if the summed GHQ-12 score is equal to 3 or higher, as our primary indicator of mental health. However, we did

not validate this threshold score by means of in-depth psychiatric interviews. We therefore confirm the robustness of our main results using summed GHQ-12 scores. We also use another screening instrument called the Mental Health Inventory 5 (MHI-5) to further examine the robustness of our results. The MHI-5 is a more focused instrument that contains five Likert items aimed specifically at depression and anxiety (e.g. Cuijpers et. al. 2009). Appendix 1 lists each of these five items.

3. Research Setting and Design[8]

3.1 Location

Malawi, the setting for this research project, is a small and poor country in southern Africa. 81% of its population of 15.3 million lived in rural areas in 2009, with most people relying on subsistence farming. The country is poor even by African standards: Malawi's 2008 GNI per capita figure of $760 (PPP, current international $) is less than 40% of the sub-Saharan African average of $1,973 (World Development Indicators Database, 2010). According to the same data source, net secondary school enrollment is very low at 24%.

3.2 Sampling

Zomba District in the Southern region was chosen as the site for this study. Zomba District is divided into 550 enumeration areas (EAs), which are defined by the National Statistical Office of Malawi and contain an average of 250 households spanning several villages. Fifty of these EAs lie in Zomba city, while the rest are in seven traditional authorities. Prior to the start of the experiment, 176 EAs were selected from three different strata: Zomba city (urban, 29 EAs), near rural (within a

[8] This section draws heavily from Section 2 in Baird, McIntosh, and Özler (forthcoming), which provides more detail on the study design and the intervention.

16 KM radius of Zomba city, 119 EAs), and far rural (28 EAs). The choice of a 16 KM radius around Zomba city was arbitrary and based mainly on a consideration of transport costs.

In these 176 EAs, each dwelling was visited to obtain a full listing of never-married females, aged 13-22.[9] The target population was then stratified into two main groups: those who were out of school at baseline (*baseline dropouts*) and those who were in school at baseline (*baseline schoolgirls*). In each selected EA, all eligible *baseline dropouts* and a percentage of all eligible *baseline schoolgirls* were sampled. The sampling percentage depended on the strata and age of the *baseline schoolgirls* and varied between 14% and 45% in urban areas and 70% to 100% in rural areas. This procedure resulted in a sample size of 3,796 young women with an average of 5.1 dropouts and 16.7 schoolgirls per EA.

3.3 Research Design and Intervention

Treatment status was assigned at the EA level and the sample of 176 EAs was randomly divided into two equally sized groups: treatment and control. In the 88 treatment EAs, all *baseline dropouts* were offered conditional cash transfers. The 88 treatment EAs were then randomly assigned to one of three groups to determine the treatment status of *baseline schoolgirls*: in 46 EAs a randomly determined share of *baseline schoolgirls* received transfers *conditional* on regular school attendance (CCT arm), while in 27 EAs a randomly determined share of *baseline schoolgirls* received *unconditional* transfers (UCT arm).[10] In the remaining 15 EAs *baseline schoolgirls* received *no* transfers.[11] We compare schoolgirls who lived in a treatment EA but who were not selected into the treatment group to girls in control villages to measure spillover effects of the program. No EA in the sample had a similar cash transfer program before or during the study.

[9] The target population of 13-22 year-old, never-married females was selected for a variety of reasons. For details, we refer the reader to Baird, McIntosh, and Özler (forthcoming).
[10] The 'conditionality' experiment was not conducted among *baseline dropouts*. As the sample size for this cohort is relatively small (889 girls in 176 EAs – approximately 5 girls per EA – at baseline), dividing the treatment group into a conditional and an unconditional treatment arm would yield an experiment with low statistical power.
[11] In these 15 EAs in which schoolgirls did not receive transfers, only *baseline dropouts* received treatment.

3.3.1 CCT ARM

After the random selection of EAs and individuals into the treatment group, the local NGO retained to implement the cash transfers held meetings in each treatment EA between December 2007 and early January 2008 to invite the selected individuals to participate in the program. At these meetings, the program beneficiary and her parents/guardians were made an offer that specified the monthly transfer amounts being offered to the beneficiary and to her parents, the condition to regularly attend school, and the duration of the program.[12] It was possible for more than one eligible girl from a household to participate in the program.

The offer to participate in the program consisted of a transfer to the parents, a transfer directly to the girl, and payment of school fees for girls attending secondary school. Transfer amounts to the parents were varied randomly across EAs between $4, $6, $8, and $10 per month, so that each parent within an EA received the same offer. Within each EA, a lottery was held to determine the transfer amount to the young female program beneficiaries, which was equal to $1, $2, $3, $4, or $5 per month. The fact that the lottery was held publicly ensured that the process was transparent and helped the beneficiaries to view the offers they received as fair. Secondary school fees were paid in full directly to the schools.[13]

Monthly school attendance of all the conditional cash transfer recipients was checked and payment for the following month was withheld for any student whose attendance was below 80% of the number of days school was in session for the previous month. However, participants were never removed from the program for failing to meet the monthly 80% attendance rate, meaning that if

[12] Due to uncertainties regarding funding, the initial offers were only made for the 2008 school year. However, upon receipt of more funds for the intervention in April 2008, all the girls in the program were informed that the program would be extended to continue until the end of 2009.

[13] Primary schools are free in Malawi, but student have to pay non-negligible school fees at the secondary level. The program paid these school fees for students in the conditional treatment arm upon confirmation of enrollment for each term. Private secondary school fees were also paid up to a maximum equal to the average school fee for public secondary schools in the study sample.

they subsequently had satisfactory attendance, then their payments would resume. Offers to everyone, identical to the previous one they received and regardless of their schooling status during the first year of the program in 2008, were renewed between December 2008 and January 2009 for the second and final year of the intervention, which ended at the end of 2009.

3.3.2 UCT ARM

In the UCT arm, the offers were identical with one crucial difference: there was no requirement to attend school to receive the monthly cash transfers. Other design aspects of the intervention were kept identical so as to be able to isolate the effect of imposing a schooling conditionality on the outcomes of interest. For households with girls eligible to attend secondary schools at baseline, the total transfer amount was adjusted upwards by an amount equal to the average annual secondary school fees paid in the conditional treatment arm.[14] This additional amount ensured that the average transfer amounts offered in the CCT and UCT arms were identical and the only difference between the two groups was the "conditionality" of the transfers on satisfactory school attendance. Attendance was never checked for recipients in the unconditional arm and they received their payments by simply presenting at the transfer locations each month.

Baird, McIntosh, and Özler (forthcoming) explains that the rules of the program were well understood by the girls in the UCT arm. These girls knew that nothing was required of them to participate in the program and they were given no rules or regulations tied to the receipt of the transfers other than showing up at the pre-determined cash transfer locations. However, qualitative

[14] Because the average school fees paid in the conditional treatment arm could not be calculated until the first term fees were paid, the adjustment in the unconditional treatment arm was made starting with the second of 10 monthly payments for the 2008 school year. The average school fees paid for secondary school girls in the conditional treatment group for Term 1 (3,000 Malawian Kwacha, or approximately $20) was multiplied by three (to calculate an estimate of the mean annual school fees), divided by nine (the number of remaining payments in 2008) and added to the transfers received by households with girls eligible to attend secondary school in the UCT arm. The NGO implementing the program was instructed to make no mention of school fees but only explain to these households that they were randomly selected to receive a 'bonus.'

interviews suggest that the girls in the UCT arm knew about the CCT intervention being conducted simultaneously in other EAs and many had friends or acquaintances in the CCT arm. Through these contacts they knew that school attendance was strictly monitored in the CCT arm and that non-compliers were penalized. The impact of the UCT intervention as a 'positive income shock' should be interpreted in this context.

3.4 HOUSEHOLD SURVEY

The data used in this paper were collected in three household survey rounds. Baseline data, or Round 1, were collected between October 2007 and January 2008, before the offers to participate in the program took place. First follow-up data collection, or Round 2, was conducted approximately 12 months later – between October 2008 and February 2009. The second follow-up (Round 3) data collection was conducted between February and June 2010 – after the completion of the two-year intervention at the end of 2009 to examine the final impacts of the program. The intervention period coincided with the 2008 and 2009 school years.

The annual household survey consisted of a multi-topic questionnaire administered to the households in which the sampled respondents resided. It consisted of two parts: one that was administered to the head of the household and the other administered to the core respondent, i.e. the sampled girl from our target population. The former collected information on the household roster, dwelling characteristics, household assets and durables, shocks and consumption. The survey administered to the core respondent provides detailed information about her family background, schooling status, health, dating patterns, sexual behavior, fertility, and marriage. Our main mental health instrument, the GHQ-12, was part of the two follow-up surveys (Rounds 2 & 3), but it was not administered at baseline. The additional mental instrument, the MHI-5, was only administered during the second follow-up survey (Round 3).

4. Estimation Strategy and Results

4.1 Attrition and Balance

Table 1 investigates whether there are differences in the attrition rates between the treatment groups under investigation in this paper. Columns (1) & (2) regress a binary indicator for inclusion in the Round 1, 2, and 3 data on the main treatment indicators, respectively for the two cohorts analyzed in this paper: *baseline dropouts* and *baseline schoolgirls*. The regression results show that the CCT did not have a significant effect on the probability of being interviewed in the two follow-up rounds in either of the two cohorts. Nor were *baseline schoolgirls* who were offered a UCT or who were assigned to the spillover group (i.e. girls living in treatment EAs who did not receive treatment) more or less likely to be interviewed at follow-up survey rounds. Together, these results suggest that the findings of program impact are unlikely to be biased by differential attrition.

In all regressions described below, we include baseline values of the following variables as controls: age dummies, a dummy for mother's education beyond primary school, a dummy that indicates if a girls lives with her mother, a dummy for having been ill during the two weeks prior to the interview, a dummy variable for ever having had sex, and an indicator for the highest grade attended. These variables were chosen because they are predictive of mental health and, as a result, improve the precision of the impact estimates.[15] Following Bruhn and McKenzie (2009) we also include indicators for the strata used to perform block randomization – Zomba Town, rural within sixteen kilometers of the town, and rural beyond sixteen kilometers.

[15] To select this set of baseline covariates, we ran a regression of mental health status among the control group in Round 2 and Round 3 on a set of baseline characteristics suggested by our review of the literature. The other baseline characteristics included in the regression were: a dummy for father's education beyond primary school, a dummy that indicates if a girls lives with her father, and an asset index (the first principal component of 25 durable goods). None of these characteristics correlates significantly with the mental health statuses of respondents in the control group in Round 2 or Round 3.

Table 2 provides a summary of these baseline characteristics. Columns (1) and (3) show the mean values in the control group for the two cohorts examined in this study. We note that *baseline dropouts* are older, more likely to be orphaned, poorer, and much more likely to be sexually active than *baseline schoolgirls*. Columns (2), (4), and (5), investigate whether the average baseline characteristics are significantly different in the treatment arms than the control group. As one would expect under a successfully implemented randomization procedure, the baseline covariates are uncorrelated with treatment status with the exception of age and highest grade attended at baseline – two variables that are highly correlated with each other. Overall, however, the experiment appears to be well balanced across baseline characteristics that are prognostic of mental health status.

4.2 SPECIFICATION

We analyze the intention-to-treat (ITT) effects of the intervention on Round 2 and Round 3 mental health indicators using cross-sectional regressions. The regression-adjusted ITT impact of the program is estimated using (straightforward variations of) the following reduced-form linear probability model:

$$Y_i = T_i^c \gamma^c + T_i^u \gamma^u + X_i \beta + \varepsilon_i,$$

where Y_i is the mental health outcome for individual i in Round 2 or Round 3. X_i is a vector that contains the baseline covariates discussed above and presented in Table 2. T_i^c (T_i^u) is a dummy variable that takes the value 1 if a girl was in the CCT (UCT) arm. The standard errors ε_i are clustered at the EA level, which accounts both for the design effect of our EA-level treatment and for the heteroskedasticity inherent in the linear probability model. Age- and stratum-specific sampling weights are used to make the results representative of the target population in the study area. To make the results comparable across survey rounds, the analysis includes respondents if and only if they were interviewed in all three survey rounds.

4.3 RESULTS

4.3.1 BASIC TREATMENT IMPACTS ON MENTAL HEALTH

Average ITT Effects

In Table 3, we present the average impact of the CCT and UCT treatment arms on the GHQ-12 binary indicator of psychological distress. Panels A and B present program impacts among *baseline dropouts* and *baseline schoolgirls*, respectively. The first two columns investigate psychological distress in Round 2 (i.e. measured during the first follow-up survey while the intervention was still in progress) and the next two columns investigate the same outcome in Round 3 (i.e. measured during the second follow-up survey shortly after the intervention had ended). Regression models presented in columns (1) and (3) do not include any controls other than treatment status, while regressions presented in columns (2) and (4) include the set of baseline characteristics described in Section 4.1 as controls.

Panels A & B show that the GHQ-12 binary indicator of psychological distress classifies approximately 45% of *baseline dropouts* and 37% of *baseline schoolgirls* in the control group as (potentially) suffering from psychological distress in Round 2.[16] Panel A shows that CCTs did not affect psychological distress among *baseline dropouts* in either round. Panel B, on the other hand, shows that CCTs did have a significant impact on the psychological distress of *baseline schoolgirls* in Round 2, for whom the probability of suffering from psychological distress was 6 to 8 pp (or at least 17%) lower.

[16] The higher rates of psychological distress among *baseline dropouts* compared to *baseline schoolgirls* are consistent with poverty and social disadvantage as important predictors of psychological distress (as *baseline dropouts* tend to come from poorer socioeconomic backgrounds). However, they could also be the result of a higher likelihood to drop out of school when suffering from a psychological distress.

The impact on psychological distress was even more pronounced in the UCT arm in Round 2.[17] *Baseline schoolgirls* in the UCT arm exhibit a 14 pp reduction in psychological distress (Panel B, columns (1) & (2)) – an improvement of 38% over the control group.[18] The F-tests at the bottom row of the table show that the large difference between the CCT and UCT arms is statistically significant at the 5% level, but only when baseline controls are included.

The results presented in columns (3) & (4) of Table 3 suggest that the large impact of the intervention on psychological distress among *baseline schoolgirls* dissipated shortly after the intervention ended. Program impacts are small and statistically insignificant in either treatment arm in Round 3. Finally we note that using baseline characteristics as controls in the impact regressions makes little difference to the findings. In the rest of the paper, we present impact regressions with baseline controls only and omit those without controls for brevity.

Robustness to the Chosen Threshold Score

Our binary indicator of psychological distress uses a threshold score of 3 in the overall GHQ-12 score to identify whether an individual is potentially suffering from psychological distress. However, as discussed in section 2, this threshold score is not universal and has not been validated in the context of Malawi. We therefore check whether the results presented in Table 3 are sensitive to the chosen threshold score. First, following Stillman, McKenzie, and Gibson (2009), Figures 1a and 1b display cumulative density functions (CDFs) by treatment status among *baseline dropouts* in Rounds 2 and 3, respectively. The threshold score we use to define psychological distress is

[17] As explained in section 4.3 the conditionality experiment was only conducted among *baseline schoolgirls* and hence our analysis of program impacts in the unconditional and conditional treatment arms is restricted to this cohort.

[18] While the reduction in psychological distress is substantial, it does not appear to be out of line with changes in mental health found in previous research. Friedman and Thomas (2008), for instance, show that the Indonesian economic crisis of the 1990s more than doubled their panel measure of the prevalence of sadness and anxiety among women – from 19.9% to 42.2%. Similarly, depending on the specification used, Stillman, McKenzie and Gibson (2009) show that migration from Tonga to New Zealand lowered their measure of the prevalence of poor mental health among women from 41.6% to a very low 3.5%.

indicated with a vertical line.[19] We observe that the CDFs for treatment and control are virtually identical in Round 2 and there is no evidence of a treatment effect anywhere along the range of the GHQ 0/12 score. In Round 3, the CDF for the treatment group even dips below that of control girls along part of the 0 to 12 interval. We conclude that the lack of evidence for a beneficial ITT effect among *baseline dropouts* is not due to the specific threshold that was chosen here.

Figures 2a and 2b display CDFs by treatment status among *baseline schoolgirls* in Rounds 2 and 3, respectively. Figure 2a shows that, in Round 2, the CDF for the UCT arm lies above that for the CCT arm up to a cumulative GHQ-12 score of 10, at which point it reaches a value of 0.99. The CDF for the CCT arm, in turn, lies above that for the control group over the entire range of GHQ-12 scores. These graphs confirm that the large ITT effects observed among *baseline schoolgirls* in Round 2 hold over a very large range of the mental health scores. The ITT effect would therefore be qualitatively similar (although not necessarily statistically significant) if a different threshold score were used. Figure 2b shows that the CDF for the UCT arm no longer lies above that for the CCT arm in Round 3. However, the CDFs for *baseline schoolgirls* in both treatment arms do still lie above that for the control group over the entire range of GHQ-12 scores. This last finding suggests that there may still be a modest program impact on mental health in Round 3 that could not be detected by the binary measure of psychological distress used in Table 3.

Table 4 complements Figures 1 & 2 using total GHQ-12 and MHI-5 scores as dependent variables. We linearly transformed these dependent variables such that they take values from 0 to 100 and are thus easier to compare. We see that using these alternative measures to a binary measure of psychological distress does not alter the main results. We again find that CCTs do not improve mental health among *baseline dropouts* (Panel A). Among *baseline schoolgirls,* we find that mental health

[19] The threshold line is drawn at a score of 2, because CDFs display the fraction of respondents with a score *lower* than, or equal to the threshold instead of *higher* than or equal to the threshold. A threshold score of 3 or higher is thus equal to a score of 2 or lower in the presented figures.

is significantly improved in Round 2 in both treatment arms (Panel B). As previously, the effect in the UCT arm is significantly stronger than that in the CCT arm. Together, the graphical analysis and the alternative outcome measures show that our results are robust to the use of alternative indicators of psychological distress.

Dimensions of Mental Health Affected

In Table 5 we investigate the impact of the intervention on the three dimensions of mental health measured by the GHQ-12 (Graetz, 1990): anxiety and depression, social dysfunction, and loss of confidence. For this analysis, we again linearly rescaled the three dimensions such that they take values from 0 to 100 so that the magnitude of the impact on each dimension is comparable to the others. Panel A shows that among *baseline dropouts* the CCT intervention did not affect any of the three components of mental health. Among *baseline schoolgirls* (Panel B), on the other hand, we do find a substantial impact of the intervention on two of the components. First, there is a significant 3 percentage point improvement in the Round 2 social dysfunction scale among respondents in the CCT arm, which appears to have persisted after the end of the intervention. Second, both the anxiety and depression scale and the social dysfunction scale improved by more than 8 pp among respondents in the UCT arm in Round 2. The impact on the anxiety and depression scale had weakened (but was still statistically significant at the 90% confidence level) by Round 3, whereas the program effect on social dysfunction had entirely dissipated.

4.3.2 Heterogeneity of Program Impacts

As we have shown above, while cash transfers were beneficial in improving mental health among *baseline schoolgirls* in both treatment arms, the program impact was significantly larger in the UCT arm than in the CCT arm. In this sub-section, we examine the elasticity of mental health

outcomes with respect to the randomized benefit levels offered separately to the adolescent girl and her parents/guardians, which provides an explanation for the heterogeneous impacts of CCTs and UCTs on mental health among school-age girls.

As described earlier the transfers to the household were split between the adolescent girl and her parents/guardians. The transfer size to the girl was randomized at the *individual level* to take an integer value between \$1 and \$5 per month, while the transfer size to the parents was randomized at the *EA level* and took the value of either \$4, \$6, \$8, or \$10 per month. In this sub-section, we re-estimate the treatment effect of the program by including additional regressors for the individual and household amounts in our analysis:

$$Y_i = I_i^C \delta^C + I_i^U \delta^U + H_i^C \phi^C + H_i^U \phi^U + T_i^C \gamma^C + T_i^U \gamma^U + X_i \beta + \varepsilon_i$$

T_i^C and T_i^U are again dummy variables for the CCT and UCT offers, I_i^C, I_i^U and H_i^C, H_i^U give the individual and household transfer amounts in each treatment arm, defined in differences from the lowest amount offered in the treatment arms (\$1 for the individual transfer, \$4 for the household transfer). The estimates of δ and ϕ thus give the marginal effect of increasing individual and household amounts by \$1 under each treatment arm, while the estimates of γ measure the impact of each treatment at the lowest total transfer amount. The standard errors ε_i are again clustered at the EA level.

The results are displayed in Table 6. Each additional dollar transferred to the parents of a *baseline schoolgirl conditional* on her school attendance increases the likelihood of her suffering from psychological distress during the intervention by 3 pp (column (3)). This detrimental effect of additional conditional dollars to the parents disappears completely after the end of the intervention (column (4)). In contrast, we detect no relationship between the parental transfer size and the mental

health of the adolescent *baseline schoolgirls* if the transfers were offered unconditionally.[20] Interestingly, column (1) shows that increased *conditional* cash transfers to the parents had a similarly detrimental effect on the mental health of *baseline dropouts* during the intervention (p-value=0.128), which also disappeared by Round 3 (column (2)). In contrast to the transfers given to the parents, transfers given to the girls seem to have mildly beneficial effects on their mental health, especially after the end of the intervention (column (4)).

These findings indicate that when the amount transferred to the parents of *baseline schoolgirls* is at its minimum value of $4/month (and $1/month to the girls), the beneficial impact of CCTs on the mental health of adolescent girls is large and roughly equal to that of UCTs. However, simply doubling the amount offered to the parents (from $4 to $8/month) in the CCT arm is sufficient to entirely wipe out this beneficial effect, while the same increase in benefit levels to the parents would not significantly affect the mental health of adolescents in the UCT arm. Cash transfers tied to an adolescent girl make her a breadwinner for her household regardless of the conditionality. Our results suggest that when the transfers become an important source of income for the entire family *and* depend on her actions each month, they might turn into too heavy a burden for her to shoulder and become detrimental to her mental health.

The beneficial program impacts on mental health are concentrated among *baseline schoolgirls*: *baseline dropouts* saw no change in their mental health outcomes as a result of being offered CCTs. The heterogeneity in program impacts between these two cohorts of adolescent girls remains even when we restrict our sample to only those who received CCT offers, which raises the question of why the impact of the same treatment would vary substantially across these two cohorts. One

[20] The F-statistics at the bottom of the table show that we cannot state with statistical confidence that the marginal impact of each additional dollar offered to the parents of a girl in the CCT arm is higher than that for a girl in the UCT arm: the p-values for this test are over 0.2 for both indicators of mental health. As the parental transfer amounts were randomized at the community level, the power of this particular statistical test is low.

plausible explanation is that the psychological burden of returning to school for *baseline dropouts* is larger than the burden of staying in school for *baseline schoolgirls*. Returning to school can imply a significant change in lifestyle and daily activities of an adolescent girl. Our data indicate that program beneficiaries who had already dropped out of school at baseline were not only spending more time in school during the program instead of doing chores or working, but they also spent significantly less time sleeping or enjoying leisure.[21] In contrast, *baseline schoolgirls* in the CCT arm have no change in their daily activities. Hence, it is possible that a significant change in time use among *baseline dropouts* associated with returning to school may have contributed to the finding of heterogeneity of program impacts by schooling status at baseline. The results are also consistent with the hypothesis that adolescent girls from more favorable socioeconomic backgrounds may experience stronger treatment effects. However, when we searched for evidence of heterogeneity of the treatment effects along the baseline covariates included in the regression analysis, as well as the baseline propensity to drop out of school and a baseline wealth index, we found no systematic evidence suggesting that the treatment effect is heterogeneous along these dimensions.[22]

4.3.3 SPILLOVER EFFECTS

There are reasons to expect that the program can also affect the mental health of girls who were not selected into treatment. The mental health of eligible non-beneficiaries living in treatment communities may, for example, deteriorate due to the increased inequality in income and educational attainment. It is also possible that program beneficiaries change their behavior towards non-beneficiaries, leading to a change in friend networks among female adolescents and possibly a feeling of social exclusion among non-beneficiaries. For siblings of treated girls, on the other hand, the mental health effects could run in the other direction, as they may indirectly benefit from the

[21] These results, not shown here, are available from the authors upon request.
[22] These findings are also available from the authors upon request.

additional income coming into their households or simply benefit from living with sisters with improved mental health. We conclude this section by examining the spillover effects of the program on untreated adolescent girls in treatment EAs.

As the percentage of *baseline schoolgirls* treated in each treatment EA was randomly determined, we can identify the causal impact of the program on eligible non-beneficiaries in treatment EAs by comparing their mental health outcomes to those in control villages. Furthermore, because the program was assigned at the individual (and not the household) level, it is possible to examine these spillover effects separately for those who live in the same household with a program beneficiary and for others. Of the 588 *baseline schoolgirls* in our study sample who were untreated in treatment villages, 71 of them lived in the same household with a treatment girl at baseline.[23]

The first column of Table 7 shows that untreated girls in treatment villages are significantly more likely to suffer from psychological distress than the control group during the intervention period. Psychological distress among this group is 6.6 pp higher than in the control group, statistically significant at the 95% confidence level (column (1)). This result points to detrimental spillover effects within treatment EAs. Columns (3) & (5) of Table 7 separately examine the spillover effects during the intervention for untreated girls who did not live in a treatment household at baseline and for those who did.[24] The results are striking: among *baseline schoolgirls* who did not live in a household with an eligible sibling at baseline, the likelihood of psychological distress in Round 2 is 10.1 pp higher in treatment EAs than control EAs. Siblings of program beneficiaries, on the other hand, seem to have reaped substantial benefits from the program: the likelihood of suffering from psychological distress in Round 2 is 8.2 pp lower in the treatment group (p-value=0.115). This latter

[23] We refer to these 71 girls as siblings even though we don't know their exact relationship to the core respondent. They are more properly defined as eligible girls in the study sample, who lived in the same household with a program beneficiary at baseline.

[24] Households containing a single adolescent girl eligible for the program may differ from those with more than one eligible girl. Therefore, to estimate unbiased impacts of spillovers for these two types of households, we control for the number of eligible girls in regression specifications presented in columns (3)-(6).

impact is large but not statistically significant at the 90% confidence level, which is not surprising given the small number of girls living with program beneficiaries at baseline. All of these spillover effects disappeared quickly after the intervention ended (columns (2), (4), & (6)).

These findings are important. First, they suggest that the program caused a large inequality in mental health between beneficiaries and non-beneficiaries in treatment areas: the increase in the probability of suffering from psychological distress among eligible non-beneficiaries not living with a treated girl is almost as high as the decrease in the same probability among program beneficiaries. Second, the findings imply that the program may have beneficial spillover effects on adolescent girls living in treated households even if they are not direct beneficiaries of the program. These girls seem to have benefited from the additional income flowing into their household (or experienced strong peer effects) without the added responsibility for the transfers to the household.

4.3.4 CHANNELS

The presented results show that the cash transfer intervention caused a marked improvement in mental health among *baseline schoolgirls*. Most of these effects had either dissipated altogether or weakened considerably quickly after the intervention ended. UCTs resulted in stronger improvements in the mental health of *baseline schoolgirls* during the intervention than did CCTs. Above, we exploited the randomized benefit levels to the girls and their parents to show that higher amounts transferred to the parents are responsible for the differential impacts of the UCT and CCT arms. An important question that remains pertains to the cause of the improvements in mental health among *baseline schoolgirls*: was it increased school participation induced by the intervention or increased income available to the girls and their parents?

We approach this question non-experimentally.[25] Table 8 gives the total GHQ-12 score by treatment status and teacher-reported enrollment status in Round 2 for *baseline schoolgirls*. For each treatment arm, the top row shows the average GHQ-12 score by Round 2 enrollment status, while the bottom row shows the average Round 2 enrollment rate in that group. The first thing to note is that there is a large difference in mental health status in the control group between those enrolled in school and those who are not. This finding is consistent with two hypotheses that are observationally equivalent: (i) school enrollment has a causal beneficial effect on mental health among this cohort and (ii) those with poor mental health select themselves out of school. If hypothesis (i) is correct, then improved school participation induced by the intervention could be responsible for the improved Round 2 mental health of *baseline schoolgirls*.

Table 8, however, suggests that even if we assume that school participation has a causal effect on mental health, this channel cannot explain the large gains in mental health of *baseline schoolgirls* observed in this study. First, the row percentages for each treatment arm show that while the CCT arm caused a substantial decline in the dropout rate of 6.4 pp (or approximately 40%, column (1)), this marginal improvement was accompanied by only a modest drop in average GHQ-12 scores from 2.52 to 2.14 (column (3)). The UCT arm, on the other hand, caused a smaller decline of 2.5 pp in the dropout rate, but in a much more pronounced improvement in average GHQ-12 mental health scores from 2.52 to 1.73. This finding indicates that, regardless of the impact of school participation, increased income explains an important part of the mental health effects observed in this study.

Second, a quick back of the envelope calculation shows that the group of girls induced to stay in school by the CCT intervention is not big enough to be responsible for the strong overall

[25] Baird, McIntosh, and Özler (forthcoming) present a similar analysis for teenage marriage and provide an informal framework to help decipher the differential impacts of CCTs and UCTs on outcomes other than schooling.

impact of the CCT arm on mental health. Assume that the average difference between the mental health of enrolled and not enrolled girls in the control group of 1.08 points is due to school participation and that the girls who are induced to stay in school would see their mental health improve by an equal amount. Even in this scenario, the effect of improved school participation represents less than 20% of the total impact of the CCT intervention on mental health.[26] Given that an overwhelming majority (more than 85%) of baseline schoolgirls was still enrolled in school at the end of the 2008 school year, beneficial program effects within this group is a more plausible explanation for the strong overall impact of both CCTs and UCTs on mental health. GHQ-12 scores among this group were 2.36 in the control group, compared with 2.00 in the CCT arm and 1.60 in the UCT arm. Hence, while it is possible that school participation causally affects mental health, we conclude that increased school participation induced by the intervention does not appear to be the main channel of the mental health improvements observed in this study.

Finally, we also see a large improvement in the mental health status of those who have dropped out of school by Term 3 if they were offered UCTs, but virtually no change if they were offered CCTs (column (2)). This finding is consistent with the fact that girls who drop out of school in the CCT arm no longer receive any transfer payments, while those in the UCT arm do. This result confirms the beneficial effect of monthly income transfers (with no strings attached) on the mental health of adolescent girls regardless of their schooling status.

5. CONCLUSION

Mental health problems are common among adolescent girls. Moreover, among this demographic group mental health problems are not only of intrinsic importance, but can also have

[26] To see this, note that the additional girls kept in school by the CCT lower the average total GHQ-12 score in the CCT group by 0.069 points (0.064*(3.44-2.36)). The CCT lowers the total GHQ-12 score in the CCT by 0.39 points (2.52-2.14). Hence, the school participation effect constitutes less one fifth of the overall effect.

severe long-run health and socioeconomic consequences. They should therefore be considered as highly relevant indicators for the wellbeing of adolescents. This paper shows that income shocks caused by a randomized cash transfer experiment had a marked impact on the mental health of *baseline schoolgirls* in Malawi. In this cohort, those who were offered unconditional cash transfers experienced a 38% reduction in the probability of suffering from psychological distress. The same group of adolescent girls who were offered cash transfers conditional on regular school attendance also experienced a significant improvement in mental health, but this impact, at 17%, was significantly smaller than that for girls in the UCT arm. Increased school participation does not appear to be the main channel through which the intervention affects the mental health of *baseline schoolgirls*. The latter finding is in line with the fact that the intervention did not affect the mental health of girls who had already dropped out of school before the start of the intervention, even though their school (re)enrollment rates increased dramatically.[27] We also find evidence that the intervention had detrimental spillover effects on eligible non-beneficiaries in treatment villages, while it may have had beneficial spillover effects on the mental health of the female siblings of program beneficiaries. To a large extent, the effects on mental health dissipated soon after the intervention ended.

The large program impact we find in the unconditional treatment arm suggests that positive income shocks can substantially improve the mental health of school-age girls.[28] However, the impacts depend on the design of the intervention. In particular, higher conditional transfer amounts offered directly to the parents tampered the beneficial impact of the program on psychological distress among *baseline schoolgirls* as well as *baseline dropouts*. When an important source of income for

[27] This finding on the program impact on school re-enrollment among *baseline dropouts* is documented in Baird et. al. (2010).

[28] Cross-sectional regressions of mental health on baseline characteristics in the control group, on the other hand, did not produce any evidence of a relationship between household assets and mental health. This finding suggests that inferences using the Engel curve can be misleading and may not reflect accurately the relationship between changes in income and mental health.

the family depends on the actions of the adolescent girl, it might place a heavy burden on her and to cause adverse effects on her mental health.

Three potential concerns are worth addressing. First, the UCT experiment was run simultaneously with the CCT experiment. Although the two experiments were run in different EAs, beneficiaries in the UCT arm were aware of the schooling requirements in the CCT arm. The impact of the UCT arm as a 'positive income shock' should be interpreted in this context.

A second concern might be that we are detecting a change in the reference point for mental health among program beneficiaries in Round 2 rather than an actual improvement in mental health. We see no compelling reason for this, since it would imply that the reference point (the "usual" state) of program beneficiaries was lower at the first follow-up compared with the control group. If anything, we would expect the reference point to have *improved* for girls in the treatment group. Moreover, if the reference point had indeed changed, then the fact that the observed treatment effects did not persist in Round 3 implies that the reference point shifted back after the intervention ended. Again, this shift would have been in the opposite direction of what we would expect. This seems like an improbable sequence of events.

A third concern may be that, while the cash transfer intervention was in progress, beneficiaries found it (socially) desirable to answer our mental health questions more optimistically than the control group. Again, we have little reason to believe that our results are driven by this type of behavior. The items of the GHQ-12 are phrased such that there are no obviously correct or incorrect answers, so it is unclear what the socially desirable answer would be. Nor was there any reason for treated girls to expect that truthfully answering the mental health questions would have had any unwelcome consequences.

Overall, the results presented in this paper indicate that mental health among adolescent girls can substantially improve when they experience positive income shocks. However, if these income shocks are administered as part of a cash transfer intervention such as the one examined here, these mental health benefits can also be quickly eroded if sufficiently large payments are made to the parents conditional on the actions of their adolescent daughters. Policy-makers would be wise to take into consideration the spillover effects as well as the heterogeneous program effects on mental health with regards to the conditionality, the transfer size, and the recipient of the transfers within the household while designing cash transfer programs.

GEORGE WASHINGTON UNIVERSITY

TINBERGEN INSTITUTE AND VU UNIVERSITY AMSTERDAM

THE WORLD BANK

APPENDIX 1: THE GHQ-12 AND THE MHI-5

The GHQ-12 consists of 12 questions. The questions are displayed below, ordered by the components they measure according to Graetz (1990).

Component 1 "anxiety and depression":

1. During the past two weeks, have you lost much sleep over worry?
2. During the past two weeks, have you felt constantly under strain?
3. During the past two weeks, have you felt that you couldn't overcome your difficulties?
4. During the past two weeks, have you been feeling unhappy and depressed?

Component 2, "social dysfunction":

1. During the past two weeks, have you been able to concentrate on whatever you are doing?
2. During the past two weeks, have you felt that you were playing a useful part in things?
3. During the past two weeks, have you felt capable about making decisions?
4. During the past two weeks, have you been able to enjoy your normal day-to-day activities?
5. During the past two weeks, have you been able to face up to your problems?
6. During the past two weeks, have you been feeling reasonably happy, all things considered?

Component 3, "loss of confidence":

1. During the past two weeks, have you been thinking of yourself as a worthless person?
2. During the past two weeks, have you been losing confidence in yourself?

Each question has five possible answers: (i) much more than usual, (ii) more than usual, (iii) same as usual, (iv) less than usual, and (v) much less than usual. Respondents who report an item as applying to them "more than usual" or "much more than usual" ("less than usual" or "much less than usual" for positively phrased items) score 1 on this item, whereas others score 0. These binary scores are summed up and summed scores of 3 or higher are classified as cases suffering from psychological distress.

The MHI-5 consists of the following 5 questions:

1. How much of the time during the past month have you been a very nervous person?
2. How much of the time during the past month have you felt calm and peaceful?
3. How much of the time during the past month have you felt downhearted and blue?
4. How much of the time during the past month have you been a happy person?
5. How much of the time during the past month have you felt so down in the dumps that nothing could cheer you up?

Each question has six possible answers: (i) all of the time, (ii) most of the time, (iii) a good bit of the time, (iv) some of the time, (v) little of the time, and (vi) none of the time. Respondents who report an item as applying to them "all of the time" score 5 and respondents who report an item as applying to them "none of the time" score 0 (vice versa for positively phrased items). Summed scores provide an indication of psychological distress.

REFERENCES

Baird, S., C. McIntosh, E. Chirwa, and B. Özler (2010). "The Short-Term Impacts of a Schooling Conditional Cash Transfer Program on the Sexual Behavior of Young Women," *Health Economics*, 19(S1), pp. 55-68.

Baird, S., C. McIntosh, and B. Özler (2011). "Cash or Condition? Evidence from a Randomized Cash Transfer Program." Forthcoming, *Quarterly Journal of Economics*, 126(4).

Bruhn, M. and D. McKenzie (2009). "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1 (4), 200-232.

Cuijpers, P., N. Smits, T. Donker, M. ten Have, and R. de Graaf (2009). "Screening for Mood and Anxiety Disorders with the Five-Item, the Three-Item, and the Two-Item Mental Health Inventory." *Psychiatry Research*, 168 (3), 250-255.

Currie, J. and M. Stabile (2006). "Child Mental Health and Human Capital Accumulation: The Case of ADHD." *Journal of Health Economics*, 25 (6), 1094-1118.

Das, J., Q. Do, J. Friedman, and D. McKenzie (2008). "Mental Health Patterns and Consequences: Results from Survey Data in Five Developing Countries." *World Bank Economic Review*, 23 (1), 31-55.

Das, J., Q. Do, J. Friedman, D. McKenzie, and K. Scott (2007). "Mental Health and Poverty in Developing Countries: Revisiting the Relationship." *Social Science and Medicine*, 65, 467-480.

Donenberg, G. R., and M. Pao (2005). "Youths and HIV/AIDS: Psychiatry's Role in a Changing Epidemic." *Journal of the American Academy of Child and Adolescent Psychiatry*, 44 (8), 728-747.

Easterlin, , R. A. (1974). "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." In: P. A., Reder, M. W. (Eds.), Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz. Academic Press, New York, 89-125.

Eisenberg, D. and E. Golberstein (2009). "Mental Health and Academic Success in College." *The B.E. Journal of Economic Analysis & Policy*, 9 (1), Article 40.

Fernald, L. C. H., P. J. Gertler, and L. M. Neufeld (2009). "10-Year Effects of Oportunidades, Mexico's Conditional Cash Transfer Programme, on Child Growth, Cognition, Language, and Behaviour: A Longtitudinal Study."*Lancet*, 374, 1997-2005.

Fernald, L. C. H., M. R. Gunnar (2009). "Poverty-Alleviation Program Participation and Salivary Cortisol in Very Low-Income Children." *Social Science and Medicine*, 68, 2180-2189.

Fernald, L. C. H., X. Hou, P. J. Gertler (2008). "Oportunidades Program Participation and Body Mass Index, Blood Pressure, and Self Reported Health in Mexican Adults." *Preventing Chronic Desease*, 5 (3). http://www.cdc.gov/pcd/issues/2008/jul/07_0069.htm.

Filmer, D. and N. Schady (2009). "School Enrollment, Selection and Test Scores." World Bank Policy Research Working Paper 4998.

Fletcher, J. M. (2008). "Adolescent Depression: Diagnosis, Treatment, and Educational Attainment." *Health Economics*, 19 (7), 855-871.

Fletcher, J. M. (2010). "Adolescent Depression and Educational Attainment: Results Using Sibling Fixed Effects" *Health Economics*, 19 (7), 855-871.

Fletcher, J. M. and B. Wolfe (2008). "Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited." *Journal of Health Economics*, 27 (3), 794-800.

French, D. J. and R. J. Tait (2004). "Measurement Invariance in the General Health Questionnaire-12 in Young Australian Adults." *European Child & Adolescent Psychiatry*, 13 (1), 1-7.

Friedman, J. and D. Thomas (2008). "Psychological Health Before, During, and After an Economic Crisis: Results from Indonesia, 1993-2000." *World Bank Economic Review*, 23 (1), 57-76.

Gao, F., N. Luo, J. Thumboo, C. Fones, S. Li, and Y. Cheung (2004). "Does the 12-item General Health Questionnaire Contain Multiple Factors and Do We Need Them?" *Health and Quality of Life Outcomes*, 2, 63.

Gardner, J. and A. J. Oswald (2007). "Money and Mental Wellbeing: A Longitudinal Study of Medium-Sized Lottery Wins." *Journal of Health Economics*, 26 (1), 49-60.

Gertler, P. (2004). "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *American Economic Review*, 94 (2), 336-341.

Goldberg, D., and P. Williams (1988). *A user's guide to the General Health Questionnaire.* Windsor, UK: NFER-Nelson

Graetz, B. (1990). "Multidimensional Properties of the General Health Questionnaire." *Social Psychiatry and Psychiatric Epidemiology*, 26, 132-138.

Gureje, O. and B. Obikoya (1990). "The GHQ-12 as a Screening Tool in a Primary Care Setting." S*ocial Psychiatry and Psychiatric Epidemiology*, 25, 276-280.

Jackson, C. (2007). "The General Health Questionnaire." *Occupational Medicine*, 57 (1), 79.

Kessler, R. C., C. L. Foster, W. B. Saunders, and P. E. Stang (1995). "Social consequences of psychiatric disorders, I: Educational Attainment." *American Journal of Psychiatry*, 152, 1026-1032.

Lagarde, M., A. Haines, and N. Palmer (2007). "Conditional Cash Transfers for Improving Uptake of Health Interventions in Low- and Middle-Income Countries." *JAMA*, 298 (16), 1900-1910.

Lewinsohn. P. M., P. Rohde, and J. R. Seeley (1998). "Major Depressive Disorders in Older Adolescents: Prevalence, Risk Factors, and Clinical Implications." *Clinical Psychology Review*, 18 (7), 765-794.

Macours, K., N. Schady, and R. Vakis (2008). "Cash Transfers, Behavior Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment." World Bank Policy Research Working Paper 4759.

Miller, C., and Maxton Tsoka, "Evaluating the Mchinji Social Cash Transfer Pilot," http://www.unicef.org/socialpolicy/files/REvised_Presentation_Evaluating_the_Mchinji_Social_Cash_Transfer_Pilot_2_July_07.pdf, 2007.

Montazeri, A., A. M. Harirchi**,** M. Shariati**,** G. Garmaroudi**,** M. Ebadi, and A. Fateh (2003). "The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version." *Health and Quality of Life Outcomes*, 1, 66.

Ozer, E. J., L. C. H. Fernald, J. G. Manley, and P. J. Gertler (2009). "Effects of a Conditional Cash Transfer Program on Children's Behavior Problems." *Pediatrics*, 123 (4), e630-e637.

Pao, M., M. Lyon, L. J. D'Angelo, W. B. Schuman, T. Tipnis, and D. A. Mrazek (2001). "Psychiatric Diagnoses in Adolescents Seropositive for the Human Immunodeficiency Virus." *Archives of Pediatrics and Adolescent Medicine*, 154, 240-244.

Patel, V., A. J. Fisher, S. Hetrick, and P. McGory (2007). "Mental Health of Young People: a Global Public Health Challenge." *Lancet*, 369, 1302-1313.

Patel, V. and A. Kleinman (2003). "Poverty and Common Mental Disorders in Developing Countries." *Bulletin of the World Health Organization*, 81 (8), 609-615.

Paxson, C. and N. Schady (2007). "Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador." World Bank Policy Research Working Paper 4226.

Prince, M., V. Patel, S. Saxena, M. Maj, J. Maselko, M. R. Phillips, and A. Rahman (2007). "No Health without Mental Health." *Lancet*, 370, 859-877.

Rivera, J. A., D. Sotres-Alvarez, J. Habicht, T. Shamah, and S. Villalpando (2004). "Impact of the Mexican Program for Education, Health, and Nutrition (Progresa) on Rates of Growth and Anemia in Infants and Young Children." *JAMA*, 291 (21), 2563-2570.

Saluja, G., R. Iachan, P. C. Scheidt, M. D. Overpeck, W. Sun, and J. N. Giedd (2004). "Prevalence and Risk Factors For Major Depressive Symptoms Among Young Adolescents." *Archives of Pediatrics and Adolescent Medicine*, 164 (3), 760-765.

Saxena, S., G. Thornicroft, M. Knapp, and H. Whiteford (2007). "Resources for Mental Health: Scarcity, Inequity, and Inefficiency." *Lancet*, 370, 878-889.

Stein, M. B. and Y. M. Kean (2000). "Disability and Quality of Life in Social Phobia: Epidemiologic Findings." *American Journal of Psychiatry*, 157, 1601-1613.

Stillman, S., D. McKEnzie, and J. Gibson (2009). "Migration and Mental Health: Evidence from a Natural Experiment." *Journal of Health Economics*, 28, 677-687.

Tait, R. J., D. J. French,and G. K. Hulse (2003). "Validity and Psychometric Properties of the General Health Questionnaire-12 in Young Australian Adolescents." *Australian and New Zealand Journal of Psychiatry*." 37 (3), 374-381.

UNAIDS (2008). Report on the global AIDS epidemic. UNAIDS: Geneva. Available from: http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/.

Weich, S. and G. Lewis (1998). "Poverty, Unemployment, and Common Mental Disorders: Population Based Cohort Study." *British Medical Journal*, 317, 115-119.

Weich, S., G. Lewis, and S. P. Jenkins (2001) "Income Inequality and the Prevalence of Common Mental Disorders in Britain." *British Journal of Psychiatry*, 178, 222-227.

Wiggins, R. D., P. Schofield, A. Sacker, J. Head, and M. Bartley (2003). "Social Position and Minor Psychiatric Morbidity over Time in the British Household Panel Survey 1991-1998." *Journal of Epidemiology and Community Health*, 58, 779-787.

World Health Organization (2004). *Global Burden of Disease: 2004 Update*, Geneva, Switzerland.

Ye, S. (2009). "Factor Structure of the General Health Questionnaire (GHQ-12): The Role of Wording Effects." *Personality and Individual Differences*, 46, 197-201.

Yusoff, M. S. B., A. F. A. Rahim, and M. J. Yaacob (2009). "The Sensitivity, Specificity and Reliability of the Malay Version 12-Items General Health Questionnaire (Ghq-12) in Detecting Distressed Medical Students." *ASEAN Journal of Psychiatry*, 11 (1), 1-8.

TABLES

Table 1: Attrition. Dependent variable = 1 if interviewed during survey rounds 1, 2, and 3

	Baseline Dropouts	Baseline Schoolgirls
	(1)	(2)
Conditional treatment	0.008	0.021
	(0.028)	(0.030)
Unconditional treatment		0.030
		(0.024)
In spillover group (untreated in treatment EA)		-0.008
		(0.022)
Number of observations	889	2,907
Mean in control	0.841***	0.893***
	(0.020)	(0.011)
Prob(Conditional=Unconditional)		0.797

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Observations are weighted to make results representative of the target population in the study EAs.

Table 2: Baseline summary statistics

	Baseline dropouts		Baseline schoolgirls		
	Mean in Control	Conditional Treatment Difference	Mean in Control	Conditional Treatment Difference	Unconditional Treatment Difference
	(1)	(2)	(3)	(4)	(5)
Mother educated beyond primary school	0.291	0.038	0.329	0.067	0.042
	(0.027)	(0.037)	(0.021)	(0.056)	(0.040)
Mother lives in household	0.646	-0.062	0.650	-0.004	0.004
	(0.028)	(0.039)	(0.017)	(0.041)	(0.034)
Ill in past 2 weeks	0.399	0.023	0.426	-0.045	-0.005
	(0.026)	(0.037)	(0.022)	(0.039)	(0.048)
Never had sex	0.299	-0.013	0.797	-0.001	-0.023
	(0.029)	(0.042)	(0.016)	(0.030)	(0.034)
Highest grade attended	6.228	-0.212	7.477	-0.231	0.419**
	(0.210)	(0.319)	(0.092)	(0.184)	(0.176)
Age	17.591	-0.415	15.252	-0.299*	0.173
	(0.161)	(0.254)	(0.068)	(0.154)	(0.125)
Randomization strata					
Zomba city	0.181	-0.035	0.347	0.112	0.070
	(0.055)	(0.069)	(0.078)	(0.134)	(0.153)
Rural area within 16km radius from Zomba city	0.698	-0.028	0.563	-0.122	-0.031
	(0.062)	(0.090)	(0.074)	(0.123)	(0.148)
Rural area outside 16km radius from Zomba city	0.121	0.063	0.090	0.010	-0.039
	(0.040)	(0.071)	(0.028)	(0.050)	(0.058)
Observations	381	370	1358	470	261

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered at the EA level.
Observations are weighted to make results representative of the target population in the study EAs.
Observations include only those respondents or households who were interviewed in survey round 1, 2, and 3.
Stars on the coefficients in columns (2), (4), and (5) indicate significantly different than the relevant control group.

Table 3: Average impact of the Zomba cash transfer program on GHQ-12 binary measure of psychological distress

Panel A: Baseline Dropouts	Round 2 (During)		Round 3 (After)	
	(1)	(2)	(3)	(4)
Conditional treatment	0.011	0.005	0.026	0.022
	(0.038)	(0.038)	(0.035)	(0.036)
Number of observations	751	751	751	751
Mean in control	0.454***	0.454***	0.320***	0.320***
	(0.028)	(0.028)	(0.022)	(0.022)
Controls included	No	Yes	No	Yes

Panel B: Baseline Schoolgirls	Round 2 (During)		Round 3 (After)	
	(1)	(2)	(3)	(4)
Conditional treatment	-0.082**	-0.063**	-0.047	-0.039
	(0.034)	(0.030)	(0.049)	(0.047)
Unconditional treatment	-0.142***	-0.143***	-0.035	-0.038
	(0.042)	(0.034)	(0.047)	(0.049)
Number of observations	2,089	2,089	2,089	2,089
Mean in control	0.374***	0.374***	0.308***	0.308***
	(0.020)	(0.020)	(0.017)	(0.017)
Prob(Conditional=Unconditional)	0.197	0.029	0.845	0.988
Controls included	No	Yes	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Observations are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey rounds 1, 2, and 3. Included baseline controls are age dummies, geographical strata, and dummies for mother's education beyond primary school, girls living in a household with their mother, having been ill over the past two weeks, never having had sex, and an indicator for the highest level of education attended. P-values for equality of coefficients calculated using F-test.

Table 4: Robustness checks using total scores on GHQ-12 and MHI-5 as dependent variables

Panel A: Baseline Dropouts	GHQ-12	GHQ-12	MHI-5
	Round 2 (During)	Round 3 (After)	Round 3 (After)
	(1)	(2)	(3)
Conditional treatment	0.434	1.446	1.148
	(1.717)	(1.591)	(1.268)
Number of observations	751	751	751
Mean in control	24.300***	16.732***	24.063***
	(1.334)	(0.895)	(0.767)

Panel B: Baseline Schoolgirls	GHQ-12	GHQ-12	MHI-5
	Round 2 (During)	Round 3 (After)	Round 3 (After)
	(1)	(2)	(3)
Conditional treatment	-2.920**	-2.584	-1.116
	(1.462)	(1.654)	(1.292)
Unconditional treatment	-7.367***	-2.330	-2.248
	(1.924)	(1.801)	(1.556)
Number of observations	2,089	2,089	2,089
Mean in control	21.140***	16.615***	23.031***
	(1.031)	(0.788)	(0.598)
Prob(Conditional=Unconditional)	0.021	0.909	0.542

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Total GHQ-12 and MHI-5 scores were transformed such that they take values from 0 to 100. Observations are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey rounds 1, 2, and 3. Included baseline controls are age dummies, geographical strata, and dummies for mother's education beyond primary school, girls living in a household with their mother, having been ill over the past two weeks, never having had sex, and an indicator for the highest level of education attended. P-values for equality of coefficients calculated using F-test.

Table 5: Impact of treatment on components of psychological distress

Panel A: Baseline Dropouts	Anxiety and Depression		Social Dysfunction		Loss of Confidence	
	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)
	(1)	(2)	(3)	(4)	(5)	(6)
Conditional treatment	-0.440	2.095	1.256	1.679	-0.285	-0.551
	(2.506)	(2.167)	(1.888)	(1.875)	(1.930)	(1.478)
Number of observations	751	751	751	751	751	751
Mean in control	23.031***	14.764***	30.184***	21.435***	9.186***	6.562***
	(1.947)	(1.136)	(1.402)	(1.147)	(1.511)	(1.116)

Panel B: Baseline Schoolgirls	Anxiety and Depression		Social Dysfunction		Loss of Confidence	
	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)
	(1)	(2)	(3)	(4)	(5)	(6)
Conditional treatment	-2.793	-0.668	-3.365**	-4.970**	-1.840	0.742
	(1.814)	(2.004)	(1.572)	(1.998)	(1.514)	(1.396)
Unconditional treatment	-8.650***	-3.936*	-8.139***	-1.790	-2.484	-0.742
	(2.453)	(2.052)	(2.098)	(2.665)	(1.739)	(1.610)
Number of observations	2,089	2,089	2,089	2,089	2,089	2,089
Mean in control	18.236***	13.741***	27.650***	22.459***	7.420***	4.832***
	(1.168)	(0.901)	(1.138)	(1.097)	(0.971)	(0.747)
Prob(Conditional=Unconditional)	0.024	0.223	0.018	0.280	0.734	0.440

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Components were rescaled such that they take values from 0 to 100. Observations are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey rounds 1, 2, and 3. Included baseline controls are age dummies, geographical strata, and dummies for mother's education beyond primary school, girls living in a household with their mother, having been ill over the past two weeks, never having had sex, and an indicator for the highest level of education attended. P-values for equality of coefficients calculated using F-test.

Table 6: Impact of transferred amounts on GHQ-12 binary measure of psychological distress

	Baseline Dropouts		Baseline Schoolgirls	
	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)
	(1)	(2)	(3)	(4)
Conditional treatment	-0.044	0.032	-0.129**	0.031
	(0.065)	(0.068)	(0.052)	(0.085)
Unconditional treatment			-0.112*	0.024
			(0.060)	(0.082)
Conditional individual amounts	-0.002	-0.006	-0.011	-0.034
	(0.022)	(0.017)	(0.017)	(0.021)
Unconditional individual amounts			-0.025	-0.051**
			(0.020)	(0.025)
Conditional household amounts	0.018	0.001	0.029***	0.000
	(0.011)	(0.011)	(0.009)	(0.013)
Unconditional household amounts			0.007	0.015
			(0.015)	(0.017)
Number of observations	751	751	2,089	2,089
Mean in control	0.454***	0.320***	0.374***	0.308***
	(0.028)	(0.022)	(0.020)	(0.017)
Prob(Conditional = Unconditional)			0.828	0.952
Prob(Conditional = Unconditional) Individual amount			0.599	0.603
Prob(Conditional = Unconditional) Household amount			0.214	0.503

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Amounts are US dollars in excess of the lowest transfer amount. Observations are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey rounds 1, 2, and 3. Included baseline controls are age dummies, geographical strata, and dummies for mother's education beyond primary school, girls living in a household with their mother, having been ill over the past two weeks, never having had sex, and an indicator for the highest level of education attended. P-values for equality of coefficients calculated using F-test.

Table 7: Spillover effects on GHQ-12 binary measure of psychological distress of baseline schoolgirls

	Entire Spillover Group		Spillovers not in Treatment Household		Spillovers in Treatment Household	
	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)	Round 2 (During)	Round 3 (After)
	(1)	(2)	(3)	(4)	(5)	(6)
Spillover	0.066**	0.011	0.101***	0.006	-0.082	0.021
	(0.029)	(0.032)	(0.032)	(0.030)	(0.052)	(0.081)
R2	0.053	0.025	0.066	0.036	0.070	0.035
Number of observations	1,922	1,922	1,853	1,853	1,427	1,427
Mean in control	0.374***	0.308***	0.374***	0.308***	0.374***	0.308***
	(0.020)	(0.017)	(0.020)	(0.017)	(0.020)	(0.017)

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regressions with standard errors (in parentheses) clustered at the EA level. Observations are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey rounds 1, 2, and 3. Included baseline controls are age dummies, geographical strata, and dummies for mother's education beyond primary school, girls living in a household with their mother, having been ill over the past two weeks, never having had sex, and an indicator for the highest level of education attended. In addition, the regressions control for the number of eligible baseline dropout siblings, the number of baseline schoolgirl siblings, and sampling frame to correct for possible structural differences among households with different numbers of eligible siblings.

Table 8: Total GHQ-12 score by School Enrollment Status in Term3, 2008

	Enrolled	Not enrolled	Total
	(1)	(2)	(3)
Control	2.36	3.44	2.52
(row %)	(85.1%)	(14.9%)	(100.0%)
Conditional treatment	2.00	3.59	2.14
(row %)	(91.5%)	(8.5%)	(100.0%)
Unconditional treatment	1.60	2.61	1.73
(row %)	(87.6%)	(12.4%)	(100.0%)
Total	2.12	3.34	2.28
(row %)	(87.5%)	(12.5%)	(100.0%)

Notes: This table presents the mean GHQ-12 scores by Round 3 enrollment status in Term 1, 2010 and treatment status. For each treatment arm, the top row summarizes the GHQ-12 score by follow-up enrollment status, while the bottom row indicates the raw follow-up row percentage in each cell. Means are weighted to make them representative of the target population in the study EAs.

FIGURES

Fig 1a: CDF of mental health of baseline dropouts in Round 2 (during the intervention)

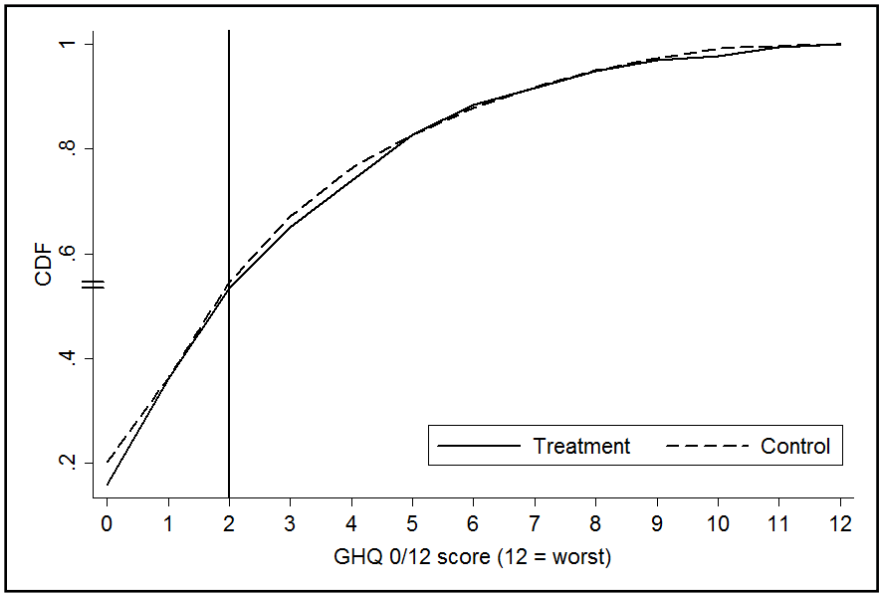
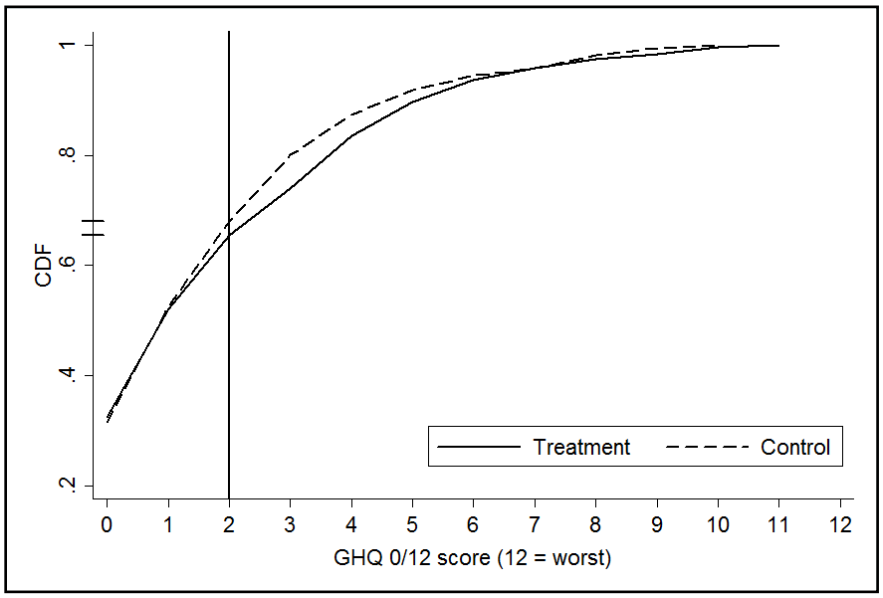


Fig 1b: CDF of mental health of baseline dropouts in Round 3 (after the intervention)



Notes: Observations in Figures 1a and 1b are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey Rounds 1, 2, and 3. Ticks on the vertical axes added where CDFs hit the threshold score. In Figure 1a the fraction of treatment and control girls below the threshold score is 0.535 and 0.546 respectively. In Figure 1b the fraction of treatment and control girls below the threshold score is 0.654 and 0.680 respectively.

Fig 2a: CDF of mental health of baseline schoolgirls in Round 2 (during the intervention)

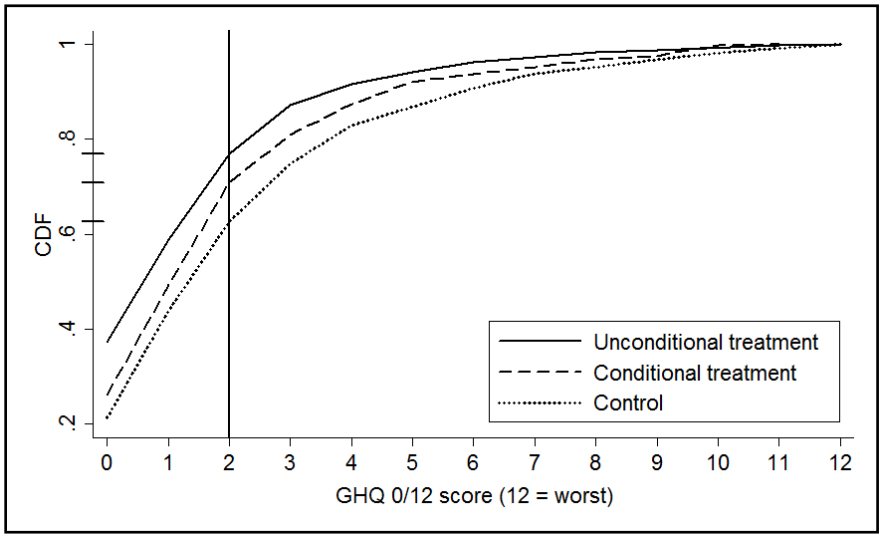
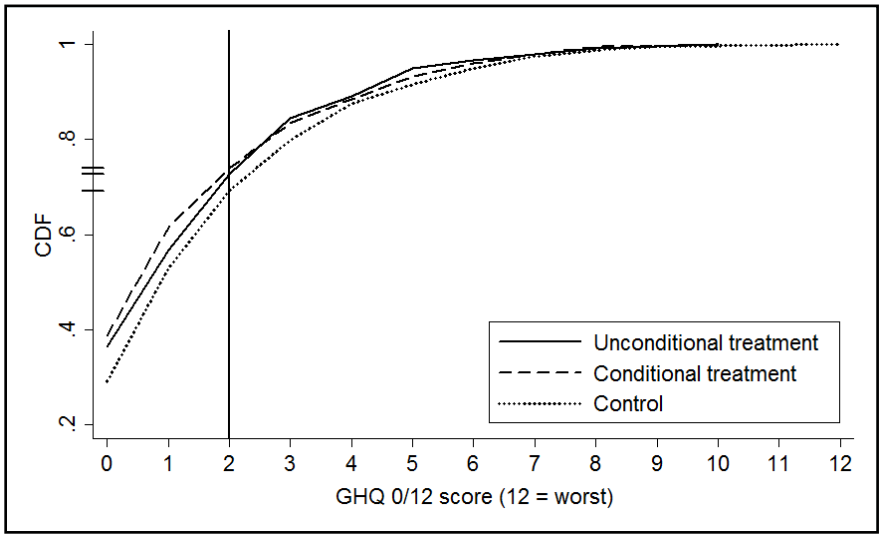


Fig 2a: CDF of mental health of baseline schoolgirls in Round 3 (after the intervention)



Notes: Observations in Figures 2a and 2b are weighted to make results representative of the target population in the study EAs. Observations include only those respondents who were interviewed in survey Rounds 1, 2, and 3. Ticks on the vertical axes added where CDFs hit the threshold score. In Figure 2a the fraction of unconditional treatment, conditional treatment, and control girls below the threshold score is 0.769, 0.709, and 0.626 respectively. In Figure 2b the fraction of unconditional treatment, conditional treatment, and control girls below the threshold score is 0.727, 0.739, and 0.692 respectively.