Direct and Interaction Effects of Cash Transfers and Psychological Interventions: Promoting Future Orientation on Economic Outcomes
Analysis Plan

Kate Orkin[*], Robert Garlick[†], Mahreen Mahmud[‡], Richard Sedlmayr[§], Johannes Haushofer[¶], Stefan Dercon[‖]

February 23, 2019

[*]University of Oxford. kate.orkin@bsg.ox.ac.uk
[†]Duke University. robert.garlick@duke.edu
[‡]University of Oxford. mahreen.mahmud@economics.ox.ac.uk
[§]University of Oxford. richard.sedlmayr@bsg.ox.ac.uk
[¶]Princeton University. haushofer@princeton.edu
[‖]University of Oxford. stefan.dercon@economics.ox.ac.uk

Contents

1 Overview

This analysis plan discusses the study "Direct and Interaction Effects of Cash Transfers and Psychological Interventions: Promoting Future Orientation on Economic Outcomes," conducted in two Kenyan counties, Homa Bay and Siaya, in Western Kenya.[1] The study is a randomised controlled trial that cross-randomises two interventions. The first is a cash transfer, administered as part of standard operations of GiveDirectly, a nonprofit organization. The other is a psychological intervention. Both interventions are randomised at village level. The study measures the separate and joint effects of these interventions.

This document describes the interventions, the core pre-registered hypotheses we commit to testing, the definitions of the outcomes of interest, and estimation and inference methods. Section 9 briefly describes several extension pieces of analysis and their relationship to the analysis in this paper.

In Appendix A, we outline a potential conceptual framework. The sketch framework informed which variables we measured, how variables are grouped together, and our strategy for accounting for testing multiple hypotheses. It also suggests some predictions about the sign of effects. However, we do not commit to using a specific theoretical model: the experiment is motivated by a broad class of models and this broad framework could be adapted or extended to help interpret particular findings. Predicted signs and magnitudes of effects may vary depending on the specific model used.

At the time of lodging this analysis plan, we have not yet completed the endline data collection.

2 Interventions

2.1 Cash transfer intervention

The cash transfer intervention is the standard programme implemented by GiveDirectly. In the areas in our study, GiveDirectly gave transfers to households fulfilling at least one of the criteria in Section 3. In cash treatment villages, GiveDirectly followed the following process:

1. GiveDirectly held a meeting open to all households in the village to explain their

[1] See https://www.socialscienceregistry.org/trials/996 for the trial registration, questionnaires and intervention materials.

programme and GiveDirectly as an organization and inform villagers that GiveDirectly would be working in their village. The eligibility criteria were not disclosed, although households were told that poorer households will be targeted.

2. One GiveDirectly team conducted a census of the village, collecting information on household names, contact information and the variables used to determine program eligibility.[2]
3. A second GiveDirectly team was given a list of eligible households for the village. They confirmed the household was eligible. If they were, they gave the household information on the programme (including the transfer size and timing and that no conditions are attached to the transfer use). This was the first time the household member(s) heard the household had been enrolled. They then registered the household for the programme, if the household consented. Transfers were offered to the household as a whole, although whichever household member is at home usually signed up to receive the transfer via M-Pesa. In roughly 86% of households in our sample, the woman is the recipient. Households were told if there is any intra-household conflict about the transfer, they can be disqualified. Households were asked to register for M-Pesa, a mobile money transfer service used to send the transfers. Registration can be done at a network of agents in most small stores. They can receive a mobile phone if they do not have one, with the cost taken off from the transfer amount.
4. All registered households were backchecked to confirm eligibility in advance of the transfers going out.[3]
5. Households were sent three mobile money transfers, made in intervals of approximately two months: a small transfer ("Token") of approximately USD100 (nominal 2016 dollars); a large transfer ("Lump Sum A") of approximately USD500; and a second large transfer ("Lump Sum B") of USD500 minus the price of the mobile phone. Transfers were typically sent at one time per month to all households scheduled to receive transfers. There is a GiveDirectly helpline that recipients can contact in case of problems.

Those trial participants who are not assigned to the cash condition receive no component of

[2]The Innovations for Poverty Action (IPA) research team conducted censuses in treatment and control villages, as described in Section 3. In all villages in Homa Bay county and roughly half of villages in Siaya county assigned to receive cash treatment, GiveDirectly conducted their own census after the research team census. In the other half of Siaya villages, GiveDirectly used the IPA census data. In villages where both censuses were collected, they produced nearly identical results for determining household eligibility: household eligibility status was the same for over 98% of households.

[3]These 'backchecks' were conducted on everyone to confirm eligibility. In addition, another audit on a sub-sample flagged for checks was conducted to confirm eligibility.

the cash transfer intervention.

2.2 Common intervention structure for psychological and placebo treatments

Trial participants were assigned to either the psychological intervention or a psychologically inactive "placebo" intervention.[4] Both interventions are structured in the same way, involving three core activities:

1. Two back-to-back ten-minute videos, viewed by participants on a tablet. A brief public service announcement, unrelated to the two videos, is shown between them.
2. A facilitated drawing exercise and discussion.
3. The distribution of a reminder calendar and stickers.

In both treatment groups, videos also provide some information in a public service announcement in the format of an advertisement break. A narrator reads words over still pictures related to the words. First, a narrator says that chlorine prevents children getting diarrhoea and notes that if there are three children who get sick from diarrhoea, using chlorine would have stopped one from getting sick. Second, a narrator says: "Only some people go to secondary school. Some young Kenyan men sometimes end up working in low-skilled jobs where they work for themselves. Others end up working in paid jobs for an employer. Those who complete secondary school are less likely to end up working independently and more likely to end up working in a paid job." We discuss how this information is used in Section 8.8.

The psychological and placebo treatments were administered by IPA enumerators. These were timed to occur in the same month as (most of) the villages in the location received the first lump sum payment of USD500. They always occurred after cash villages had received token payments. Within cash and placebo and cash and psychological intervention villages, some people got the first lump sum cash transfer before and some got it after participating in the psychological/placebo intervention.

The psychological and placebo interventions were administered differently in Homa Bay and Siaya. In Homa Bay, individuals watched the video and completed the exercises with a single facilitator (with a target time of 60 minutes) at their homes. In Siaya, people watched the video in pairs on a tablet. Groups of 2-4 people completed the exercise together, with a pair

[4]See https://www.socialscienceregistry.org/trials/991 for information on the piloting of the psychological intervention and placebo. All intervention materials including all videos are linked to on https://www.socialscienceregistry.org/trials/996.

of two facilitators (with a target time of 90 minutes). Within a village in Siaya, people were randomly invited to intervention sessions within the village (e.g. 'Tuesday afternoon') and those who arrived at the same session were randomly divided into groups. The precise group assignment protocol is described in our pre-analysis plan for data collected straight after the psychological/placebo intervention, posted in this trial registry entry.

2.3 Psychologically active intervention

The videos narrate the life stories of two role models similar to the audience who model specific behaviours and mindsets.[5] The videos describe how they came to be successful, the obstacles they faced and some lessons they drew from their experiences.

The characters are also shown doing exercises. After watching the video, the participants then go on to do these exercises themselves, with a facilitator. The exercises draw on elements from existing psychological interventions such as Best Possible Selves interventions (King, 2001; Oyserman, 2006; Peters et al., 2010), Mental Contrasting and Implementation Intentions (MCII) interventions (Duckworth et al., 2013; Oettingen and Gollwitzer, 2010), and personal goal-setting (Morisano et al., 2010; Stadler, Oettingen, and Gollwitzer, 2009).

1. Characters are depicted imagining their Best Possible Selves, envisioning themselves in a desirable future. This image is a personalised representation of their goals. Characters contrast their aspiration with aspects of their reality that impede the realisation of this future. Similarly, in the exercises after the video, participants were asked to imagine their lives in five years "after everything has gone as well as it possibly could". They drew what they imagined and explained it to the group or the fieldworker, including discussing how it was different to their lives at present.
2. Characters were shown identifying smaller, more immediate goals that would lead toward their long-term aspirations and delineated specific strategies for achieving these goals. These smaller steps were linked directly to achieving the larger long term goal. Similarly, participants worked with the fieldworker to make goals clear and specific and then to rank goals according to what they thought was both achievable and possible. They listed steps to work towards their highest ranked goal.
3. Characters were shown making specific plans for how they would adjust plans when things went wrong, using the if-then approach. They were also shown facing substantial

[5]Our thanks to Catherine Thomas and Michala Iben Riis-Vestergaard (Princeton University) and Pat Olvera and Rita Wachera (Khangarue Media) for their work on video development, scripting and production with the research team. Thanks to Carol Dweck for her comments on the script.

obstacles and planning how to overcome them. Similarly, in the exercises, after making plans towards their goals, participants worked out potential obstacles to achieving their goal and strategies they could adopt to overcome these obstacles.

In their descriptions of their experiences, the characters promote a growth mindset (Dweck, 2012; Aronson, Fried, and Good, 2002). They describe how one's skills and abilities can be changed through effort and that learning occurs throughout one's life. The characters draw on existing opportunities and resources within their immediate vicinity. They are not afraid to try new things and describe the process of addressing and overcoming challenges as enjoyable. They anticipate setbacks and obstacles, treat them as an opportunity to learn, rather than a failure, and persist despite setbacks. They also describe developing self-efficacy (Bandura, 1997). They acknowledge facing self-doubt, but describe growing in confidence and believing more in their abilities as they face and overcome challenges.

At the conclusion of the exercise, participants received a single-page calendar depicting the two role models from the videos, and sayings that described the core spirit of each video. Participants also received a set of stickers from which they could choose ones that reminded them of the exercise, and were encouraged to place these on the calendar.

2.4 Psychologically inactive (placebo) intervention

Trial participants who were not assigned to the psychological intervention were instead assigned to a placebo intervention. This followed the same format as the psychologically active intervention, including a video, an exercise and a calendar. The placebo video contains all potentially new information in the psychologically active videos. We included at least one shot of every scene and character from the other videos, including introductory shots of scenery. For example, in the psychologically active video, Josephine overfeeds her chickens so that they do not lay eggs. In the placebo, we show the same shots of a poultry house, chickens and eggs. A narrator notes that people farm chickens and that if they are overfed they will not lay eggs. However, we exclude any elements that are likely to manipulate critical psychological variables, such as characters, insightful narratives, shots of people conveying obvious emotion, or music.

The placebo group also participated in an exercise with exactly the same elements: they were reminded of the content of the video, they discussed the video and facts presented, and they drew the scenes that were most memorable. Placebo participants also received a calendar (with a sunset and acacia tree) and stickers, but were not told that the stickers represented

their goals. Placebo participants watched the videos and completed the exercises, like the psychologically active intervention, either individually in Homa Bay county or in small groups in Siaya county.

3 Sampling and treatment assignment

The sampling and treatment assignment scheme is based on household census data collected by the study team. IPA field officers completed a short census with all consenting households in all villages in target areas, which also captured information on village amenities. We used these census data to determine village and household eligibility, assign villages to treatment, and draw samples of eligible households in eligible villages before GiveDirectly entered villages.

3.1 Sample of villages and randomisation

We received a list of eligible villages from GiveDirectly, including all villages listed in the 2009 census but excluding urban or peri-urban areas. We combined some pairs of small villages.[6] We also excluded any village with fewer than 15 eligible households, defined below.

We assigned eligible villages to treatment using a sequential stratified random assignment algorithm. The algorithm was sequential because we assigned groups of villages to treatment at different times, as household census data became available. The first group of 107 villages was randomized in April 2016. The second group of 132 villages was randomized in June 2016. The third group of 132 villages was randomized in October 2016. The fourth group of 44 villages was randomized in February 2017. We stratified treatment assignment on four variables, all collected during the household census. This assignment occurred before the baseline surveys.

1. Sublocation: This is an administrative division in Kenya containing roughly 10-50 villages. We constructed sublocation blocks as pairs of geographically adjacent locations. The first, second, third, and fourth groups of villages contained respectively four, three, three, and two sublocation blocks.
2. Village amenities: We calculated the first principal component of village-level indicators equal to one if the village contains a primary school, high school, vocational school, market, and clinic (measured in the census). We then created an indicator variable

[6] We aggregated pairs of villages only if they had previously been administered as a single village, were geographically contiguous, and each contained fewer than 40 households.

equal to one if the village amenity index exceeded the sample median.[7]

3. Village assets: We calculated the first principal component of household-level indicators equal to one if the household owns a solar panel, television, fridge, iron, radio, watch or clock, telephone, bicycle, motorbike, truck, or car. We then calculated village-level averages of this index and created an indicator variable equal to one if the village asset index exceeded the sample median.
4. Village size: We calculated the number of households in each village, then created an indicator variable equal to one if the village size exceeded the sample median.

This yielded 32, 24, 24, and 8 stratification blocks in the first, second, third, and fourth groups of villages respectively. We then implemented a three-stage stratified random assignment. In the first stage, we randomly assigned villages in each stratification block to the four treatment types in groups of four. If the number of villages in any stratification block was not a multiple of four, then we proceeded to the second stage of the randomization. Here we constructed "large stratification blocks" containing leftover villages that have the same values of the sublocation, amenity, and asset variables but different values of the size variable. We randomized sets of four leftover villages within each of these large blocks. If the number of villages in any large block was not a multiple of four, we then grouped all remaining villages together and randomly assigned sets of four to treatment types. This randomization scheme prioritizes balance on sublocation, amenities, and assets ahead of balance on size.

3.2 Sample of households

We conducted our own censuses in all villages (both treatment and control), before GiveDirectly entered villages. GiveDirectly later conducted separate censuses (after our baseline) in Homa Bay and part of Siaya. Respondents' answers may differ across these censuses. We used exactly the same survey questions and criteria to define elgibility but some households flagged as "GiveDirectly-eligible" by the research team may not be regarded as eligible by GiveDirectly and vice versa. In other parts of Siaya, GiveDirectly used our census to target transfers.

In Homa Bay and parts of Siaya, we define a household as eligible to participate in the study if it satisfies at least one of the following criteria, imposed by GiveDirectly:

1. household's per capita housing space is less than 62,000cm^2
2. household has no telephone AND has a mud floor

[7] We constructed the sample median separately for the first, second, third, and fourth groups of villages.

3. household head is a widow AND has a mud floor
4. household has an orphan child
5. household is homeless

and none of the following criteria, imposed by the research team:

1. household is polygamous (due to difficulties associated with household definition)
2. household head is a child (for consent reasons)
3. household is homeless (due to difficulty finding them)
4. household does not contain an adult female (since the chosen psychological intervention is aimed at adult females)
5. household's GIS coordinates are judged to be incorrect[8]
6. household's per capita housing space is more than 58,000cm^2 (to maximise overlap with the GiveDirectly per capita housing criteria, accounting for measurement error)

In some parts of Siaya, we lowered the per capita housing space cutoff to reflect changes in GiveDirectly's targeting criteria.

In each village, we randomly drew two samples of eligible households from the census: the "target" and "reserve" households. Field officers were instructed to find each target household for the baseline survey. If a target household refused to participate or could not be located (e.g. due to migration), the field officers included one household on the reserve list as a replacement. We define the study sample as all households that completed the baseline survey. The idea of reserve households was used only for the baseline survey. In latter rounds of data collection, households that refused to participate or could not be located are included in the sample and treated as attriters.

We sampled up to 18 target and 6 reserve households in Homa Bay villages. In Siaya, where the villages are typically larger, we sampled up to 24 target and 18 reserves households in some villages and did not impose an upper limit on the number of target households in some other villages.

[8]We flagged a household as having incorrect GIS coordinates using a two-stage process. First, we calculated the median latitude and longitude within each village and flagged any household more than 3km from this joint median. Second, we calculated the mean latitude and longitude within each village excluding the flagged households, calculated the mean distance from each household to this joint mean, and flagged any household whose distance was more than 3 times the mean.

4 Data

4.1 Household survey

Our main data source is baseline and endline household surveys. Before GiveDirectly entered communities or conducted their census, we conducted a household census (as described above) and baseline household survey. Baseline surveys began in April 2016 and ended in March 2017. Cash transfers and psychological interventions began in September 2016 and ended in July 2017. Baseline surveys were always conducted before any intervention occurred in the area, but different parts of the survey team ran baseline and intervention at the same time in different areas. Endline surveys began at the end of May 2018 and will finish in February/March 2019. For a few psychological variables, we collected manipulation checks data immediately after the psychological intervention (or placebo treatment) was administered.

4.2 Price survey

At baseline, we worked with the field team to create a list of all markets in and adjacent to the study area, including information on which days these markets are open. We identified 31 markets in Homa Bay, which were typically open one or two days a week, with some open daily in the evenings. We identified 24 markets in Siaya, about half of which were open daily in the evenings, and the other half of which were typically open two days or evenings each week. We collected price information for the most commonly purchased goods and services at each market at baseline and endline, including food products, livestock prices, livestock input prices, non-food items, services and wages for different types of labour. Price surveys were carried out at the same time as the household surveys, in August 2016 in Homa Bay, and November 2016 in Siaya at baseline, and in May-June 2018 in Homa Bay and September 2018 in Siaya at endline. We also selected 5 markets in Homa Bay and 6 markets in Siaya, chosen to ensure wide geographical coverage and for logistical efficiency, for regular midline price surveys. In Homa Bay, these were surveyed in September, October, November and December 2016, as well as January, February, March, April, May, June and October 2017. In Siaya, midline surveys were carried out in December 2016 and January, February, April, May, June, September and October 2017. Village price surveys were conducted in all of the villages in our sample at endline, at the same time as the household surveys. Market price surveys were collected on paper and construction of the dataset is not yet complete.

4.3 Analysis to date

Two principal investigators and three research assistants used the early weeks of the endline data in July 2018 to ensure survey forms worked, assess data quality, shorten questionnaires and develop high frequency and consistency checks. They have also cleaned the full baseline dataset.

At the time of lodging this plan, the full endline dataset does not exist and no treatment effects have been estimated for the endline household survey datasets. Although this plan covers only analysis on people eligible for treatment, we also collect data on people both eligible and not eligible for treatment but living in treatment villages. Village and treatment identifiers and markers of whether participants are eligible for treatment are held by the survey partner, IPA, and will be provided upon lodging of this plan. Data is stored on an encrypted shared box folder with access only for specific named parties i.e. it is not available on shared file systems. IPA field managers working in Kenya have access to the full dataset, but have not estimated treatment effects for any outcomes.

We pre-registered analysis of the manipulation checks data.[9] The objective was to test if the psychological constructs that are presumed to mediate the long-term impacts were in fact successfully manipulated through the psychologically active intervention. We have examined results from this analysis and this informed the decision on whether or not to proceed with the endline survey.

We also conducted a tracking exercise with village elders before the endline survey to update village locations within the study area for people who had moved (so they were on tracking lists in the right village). We collected a separate small dataset on whether households had migrated and where to, to decide whether to track. We lodged an analysis plan to examine treatment effects on migration.[10] As there are small but statistically significant treatment effects on migration, we are tracking households which have moved.

[9]The analysis plan was lodged at https://www.socialscienceregistry.org/trials/996 on July 08, 2017.

[10]The analysis plan was lodged at https://www.socialscienceregistry.org/docs/analysisplan/1581/document on May 08, 2018.

5 Estimation and inference

5.1 Estimation

We will estimate models of the form

$$Y_{iv} = \text{Cash}_v \cdot \beta_C + \text{Psych}_v \cdot \beta_P + \text{Cash}_v \cdot \text{Psych}_v \cdot \beta_{CP} + \mathbf{X_{iv}} \cdot \mathbf{\Gamma} + \epsilon_{iv}, \tag{1}$$

where i and v index individuals and villages, Y_{iv} denotes the outcome of interest measured in the follow-up, $Cash_v$ and $Psych_v$ are indicator variables equal to one for villages assigned to receive respectively cash and psychological treatments, and $\mathbf{X_{iv}}$ is a vector of prespecified covariates: the outcome of interest measured in the baseline, Y_{0iv}; a vector of stratification block fixed effects; a vector of endline month fixed effects to account for seasonality; baseline values of household size, the asset aggregate, respondent education, age and the self-belief index; and an indicator for the endline being answered by a proxy respondent (where the respondent targeted for the psychological indicator is not available).[11,12] We chose these covariates by regressing outcome variables for part of the endline sample on various combinations of baseline covariates and selecting a combination of covariates that explained a large portion of the outcome variation. The choice of covariates is designed to improve the precision of treatment effect estimates by absorbing outcome variation. This exercise was performed before treatment assignments were observed and before endline data were available for the full sample.

Where Y_{iv} is not measured in the baseline, we usually identify a conceptually related baseline measure that is correlated with Y_{iv} and use that in place of Y_{0iv}. If no conceptually similar baseline measure is available, we analyse outcomes using Equation 1 omitting Y_{0iv}. Where Y_{iv} is measured in the baseline but is missing for some observations, we replace the missing values with the sample mean and include a missing data indicator $\mathbf{1}\{Y_{0iv} \text{ missing}\}$ as an additional regressor. If an outcome index includes both variables collected and not collected at baseline, we construct the index based on variables that were collected at baseline. If the baseline value of any other covariate in $\mathbf{X_{iv}}$ is missing, we replace the missing values with the sample mean and include a missing data indicator.

[11] We will use 'large stratification block' fixed effects, as defined in Section 3. This yields 16, 12, 12, and 4 indicators in the first, second, third, and fourth groups of villages respectively.

[12] We do not ask some modules, such as psychological measures, for proxy respondents because they were not targeted for the psychological intervention.

Outcomes in the endline survey are measured at the respondent, household, or household member level. We describe conversions from outcomes measured at the household member level to household-level outcomes in Section 7.

5.2 Inference

For each outcome, we will test three statistical hypotheses that follow directly from the conceptual framework:

(a) Assignment to the cash transfer group has no effect on the outcome relative to assignment to the cash control group, $\beta_C = 0$ ('cash effect').

(b) Assignment to the psychological intervention group has no effect on the outcome relative to assignment to the placebo intervention group, $\beta_P = 0$ ('psych effect').

(c) Assignment to the cash transfer and psychological intervention group has the same effect on the outcome as sum of the cash and psychological effects, $\beta_{CP} = 0$ ('interaction effect').

We will also test two additional hypotheses that are relevant for evaluating the two interventions but are not used to explore the underlying economic model:

(d) Assignment to the cash transfer and psychological intervention group has the same effect on the outcome as assignment to the cash transfer and placebo intervention, $\beta_P + \beta_{CP} = 0$ ('additionality effect'). This measures the effect of adding the psychological intervention to the cash transfer.

(e) Assignment to the cash transfer and psychological intervention group has the same effect on the outcome as assignment to the placebo intervention and cash control group, $\beta_C + \beta_P + \beta_{CP} = 0$ ('joint effect'). This measures the effect of both interventions relative to no interventions.

All hypothesis tests will be based on variance-covariance matrices that allow serial correlation of errors within villages and arbitrary heteroskedasticity (i.e. cluster-robust standard errors).

5.3 Adjustments for multiple testing

We group outcomes in Section 7 into families. Each family corresponds to a parameter or variable in the conceptual framework, outlined in Appendix A. We hypothesize that variables within a family are likely to respond to a treatment in similar ways. We use these families

to adjust for multiple testing in two ways.

First, we construct one scalar summary measure for each family that combines the information from all outcomes in the family (except outcomes that are explicitly listed as secondary analysis or robustness checks). We then estimate average intention-to-treat effects on each aggregate or index using equation 1. This provides an overall measure of how treatment jointly changes the outcomes in that family. The scalar summary measure for each family is either a money-metric aggregate or an inverse-covariance weighted average. The money-metric aggregates are constructed by scaling all outcomes within the family to use a common time period and adding these up. The inverse covariance-weighted averages follow Anderson (2008).[13]

Second, we also report more detailed information about how treatment changes each individual outcome in each family. Our conceptual framework predicts that individual outcomes in each family should move in the same direction. But these need not hold in extensions of the framework that allow for differences in returns or risk profiles across activities, or if our framework is altogether incorrect. We thus show effects on each individual outcome in each family. We estimate sharpened q-values that control the false discovery rate (FDR) across outcomes within each of the families (Benjamini, Krieger, and Yekutieli, 2006).[14]

We will not adjust inferences when we conduct tests of conceptually different hypotheses. For example, consumption and investment are theoretically distinct concepts that may move in different directions in response to the same intervention. Indeed, we do not expect some measures of expectations, preferences or psychological characteristics to be affected by some treatments. Hence, we will not adjust inferences for multiple testing across outcomes in different families, either the six core economic families or the families measuring mechanisms. We will not adjust inferences for multiple testing across different treatments as the cash transfer and psychological intervention are predicted to have different effects on the same outcomes.

[13]We will first re-code all primary outcomes so that higher values correspond to "better" outcomes. We will then standardize the outcomes to have mean zero and standard deviation one in the placebo intervention group. We will calculate the average of the standardized constituent outcomes, weighted by the inverse covariance matrix, and standardize this weighted average to have mean zero and standard deviation one in the placebo intervention group. We will estimate the covariance matrix and hence the weights using only observations that have non-missing values for all outcomes in the index. Where a specific outcome value is missing for a respondent, we calculate the value of the index for that respondent using the remaining outcomes.

[14]Rather than pre-specifying a single q, we report the minimum q-value at which each hypothesis is rejected. The FDR controls for the proportion of false positives, which is relevant if one is interested in the proportion of the outcomes within a family affected by treatment.

We note that best practices for multiple test adjustments are still evolving in economics, and may adjust our strategy if best practice shifts after registering this analysis plan.

5.4 Testing for treatment effect decay through time

We will use a different estimating equation to test for treatment effect decay through time for outcomes that are measured in both the manipulation checks (MC) survey and the endline (EL) survey. These outcomes are marked by ‡. We will stack observations from the two surveys and estimate the model:

$$\begin{aligned} Y_{iv} = &\text{Cash}_v \cdot \mathbf{1}\{MC = 1\}\beta_C^{MC} + \text{Cash}_v \cdot \mathbf{1}\{EL = 1\}\beta_C^{EL} \\ &+ \text{Psych}_v \cdot \mathbf{1}\{MC = 1\} \cdot \beta_P^{MC} + \text{Psych}_v \cdot \mathbf{1}\{EL = 1\} \cdot \beta_P^{EL} \\ &+ \text{Cash}_v \cdot \text{Psych}_v \cdot \mathbf{1}\{MC = 1\} \cdot \beta_{CP}^{MC} + \text{Cash}_v \cdot \text{Psych}_v \cdot \mathbf{1}\{EL = 1\} \cdot \beta_{CP}^{EL} \\ &+ \mathbf{X_{iv}} \cdot \mathbf{\Gamma} + \epsilon_{iv}, \end{aligned} \tag{2}$$

restricting the sample to respondents surveyed in both rounds. We will use village-level cluster-robust standard errors and test $\beta_j^{MC} = \beta_j^{EL}$ for $j \in \{C, P, CP\}$.

5.5 Robustness checks

We will estimate two modified version of equation 1 that use different conditioning variables. This exercise is designed to test if the estimates of $(\beta_C, \beta_P, \beta_{CP})$ are robust to conditioning on different sources of outcome heterogeneity. We will report results from this exercise only when point estimates for the three parameters of interest are substantially different to the results from estimating equation 1. The first modified version will exclude all pre-specified baseline covariates. The second modified version will also condition on:

- Fixed effects for the facilitators who conducted the psychological and placebo interventions.
- The time lag between the psychological or placebo intervention and endline dates.
- Any baseline variables that are substantially imbalanced across treatment groups.
- The relative timing of the psychological intervention and the first lump sum cash transfer.

5.6 Adjustments for attrition

The main results will be presented without adjustment for attrition (i.e. households not surveyed in the follow-up) or unit non-response (i.e. individual questions not answered in the follow-up). We will implement two analyses to characterize attrition:

1. We will compare the attrition rate by assigned treatment status. We do this by estimating model (1) using an indicator for attrition as an outcome (and omitting Y_{0iv}) and testing if any of the following linear combinations of parameters equal zero: $(\beta_C, \beta_P, \beta_{CP}, \beta_P + \beta_{CP}, \beta_P + \beta_P + \beta_{CP})$.[15]
2. We will regress an attrition indicator on a vector of baseline covariates, report the marginal effects, and test if each marginal effect is different to zero. The baseline covariates for this analysis will be respondent age, education at baseline, marital status and values of the the self-beliefs index; household assets at baseline, consumption at baseline, and baseline size; indicators for the field officers who administered the psychological intervention, sublocation indicators, and indicators for the month of the follow-up survey. We use heteroskedasticity-robust standard errors clustered by village.

If we find that the difference in attrition between any two treatment groups is larger than 2 percentage points or is statistically significantly different to zero, we will conduct two analyses to assess if our treatment effect estimates are sensitive to this attrition:

1. We will use the estimates from the previous analysis to construct the predicted probability of missing data for each observation, estimate model (1) using inverse probability weights, and implement the same hypothesis tests described in Section 5. We will construct standard errors using a two-stage bootstrap algorithm where we estimate both weights and the regression parameters in each bootstrap iteration and resample villages.
2. We will construct bounds on parameters $(\beta_C, \beta_P, \beta_{CP})$ using the trimming procedure described in Lee (2009).

[15] In the pre-analysis plan for the manipulation checks data, we specified we would test if there are any pairwise differences in attrition between groups. In this pre-analysis plan we test if attrition differs between groups used in the primary comparisons listed in Section 5.2. We will retrospectively apply this new approach to the manipulation checks data.

5.7 Heterogeneity analysis

We will estimate heterogeneous treatment effects across several baseline variables by augmenting equation 1 to include the baseline measure of interest and a vector of interactions between the vector of treatment interactions and the baseline measure of interest.

1. An indicator variable equal to one if the baseline value of the outcome is above the sample median.
2. An indicator variable equal to one if respondent's age is above the sample median.
3. An indicator variable equal to one if the respondent's education level is above the sample median.
4. An indicator variable equal to one if the respondent's asset index is above the sample median.
5. An indicator variable equal to one if the respondent is unmarried or widowed.
6. An indicator variable equal to one if the respondent's household size is above the sample median.
7. An indicator variable equal to one if the respondent is above the sample median on the index of self-beliefs.
8. An indicator variable equal to one if the respondent is above the sample median on the index of aspirations about future outcomes.
9. An indicator variable equal to one if the respondent is above the sample median on the index of expectations about future outcomes.

Both the psychological and cash treatments may shift both the mean and distribution of outcomes. For the main outcomes listed in Section 6, the index of self-beliefs, the expectations and aspirations indices (as well as other psychological mechanisms where there are any treatment effects), we will also estimate quantile treatment effects. We will use an approach adapted from Firpo (2007). First we reweight the distribution of outcomes in each treatment group using weights

$$\omega\left(g; Y_{0iv}, \alpha_v\right) = 1/Pr\left(G_v = g | Y_{0iv}, T_{iv}, \alpha_v\right) \tag{3}$$

estimated from a multinomial logistic regression of treatment group indicators on stratum fixed effects and baseline outcome values. We will winsorize the top and bottom percentiles of the weight distribution in each group and normalize the weights to sum to one. Second, we will estimate differences between the weighted outcome distributions at percentiles 5, 10, 25, 50, 75, 90, and 95 to obtain the quantile treatment effects. We may report other scalar

statistics derived from the estimated distributions to summarize the information obtained from the quantile treatment effects. We will conduct inference by bootstrapping the entire estimation process (logit regression, prediction, winsorizing, normalizing, quantile regression, summary statistic estimation) with village resampling.

6 Families of outcomes

Our conceptual framework focuses on the effects of the interventions on six core economic concepts. We view these as theoretically distinct. We discuss in the conceptual framework why, even within one treatment, effects need not go in the same direction. Each concept is captured by one of the following indices or aggregates:

1. Stock of assets: the total value of non-land household assets;
2. Revenue from economic activity: total household revenue, including from agriculture, livestock rearing and produce and non-agricultural activities, as well as household labour earnings (including casual and salaried labour and migration);
3. Investment into economic activity: total expenditure on investment into agriculture, livestock rearing and produce and all non-agricultural activities;
4. Investment into human capital: an index including education expenditure and participation;
5. Labour supply: an index capturing total household labour supply to agriculture, livestock rearing and produce, non-agricultural activities and labour outside the household (including casual and salaried labour and migration); and
6. Consumption: total household consumption expenditure excluding housing.

We also measure outcomes to understand the mechanisms behind any changes in economic outcomes, and in particular to pick up changes in decision-making or in behaviour. We examine eight groups of variables, measuring expectations, preferences and psychological characteristics. As outlined in the conceptual framework, we do not expect changes for some combinations of treatments and hypotheses, In particular, we do not think the psychological intervention will affect preferences, depression or cognitive load.

1. Expectations for one's future outcomes, an index of expectations for the future;
2. Expectations about the potential increase in income from various investments;
3. Aspirations for one's future outcomes, an index of aspirations for the future;
4. Self-beliefs, an index of psychological scales capturing self-efficacy, locus of control and growth mindset;

5. Time preferences, an index capturing present bias and discount rates;
6. Risk preferences;
7. Depression; and
8. Cognitive load.

We also examine a 'Level of information' index which summarises two alternative tests of potential mechanisms through which the psychological intervention might have effects. We do not expect differences between groups in these variables because we sought to ensure there were no differences in the information provided by the psychological and placebo interventions. The index includes measures of how well respondents remember information from the videos and whether they simply mimic activities in the video.

Finally, we conduct exploratory analysis to test whether the psychological intervention affects decisions in income-generating activities, even if households cannot afford new investment in economic activities or investments have not yet yielded returns. We examine whether households are seeking more information or investing in new technologies or production techniques, using one index of 'Technology adoption and information seeking'.

7 Outcome definitions

7.1 Definitions across sections

- Variable names indicated in parentheses are from the endline household survey questionnaire, the most recent version of which is lodged with the pre-analysis plan. Wherever possible, we construct baseline variables using the same variable; the variable labels are kept the same if the variable is the same. If baseline variables are significantly different, we note which variable is used as a control in a footnote. Outcomes marked with * have no baseline analogue.
- For Homa Bay, all agricultural data relate to the 2017 long and the 2018 short rains. For Siaya, they relate to the 2018 long and short rains.
- We list the recall period used to collect data. Where variables within an aggregate are collected using different recall periods, we will scale variables appropriately before constructing the aggregate (e.g. dividing variables collected over 12 months by 12 to create a monthly variable).
- Whenever we ask households for payments given or received, we include payments in cash or in kind and ask for estimates of the value of in kind payments.

- For all items measured in Kenyan shillings, we will not adjust for spatial price variation within a survey round (other than using prices from as local a geographic unit as possible). Analysis will be done in nominal values. This assumes that there are no differential price (or inflation) effects across treatment villages and between treatment and control villages. In companion work (see Section 9), we explicitly test for spatial and temporal price effects using the price survey data described in Section 4.2. We will construct an intertemporal deflator to express values in real terms. If the companion study finds price differentials between treatment and control villages, we will also construct a spatial deflator and use it to test for robustness of the interpretation of the treatment effects, to ensure we distinguish treatment effects in real terms from nominal effects (driven by price changes in treated villages relative to control villages). If we find no differential price effects, we will use the intertemporal price deflator to offer context on of the size effects since baseline in real terms (for example, to allow us to comment on whether values of specific stocks and flows changed since baseline in real terms).
- We include checks for most revenue, consumption and labour supply data (e.g. for sales of produce, we capture both the total units sold and the total sale value, so can use this to adjust units or sales which are very large outliers), which we to hanwill use d-clean data before receiving treatment assignments. We also winsorize variables at the top 1%.
- For psychological outcomes where we randomised the order in which questions appear in the questionnaire (psychological scales, risk and time preferences), we control for the order in which the questions appear in the main specification.

7.2 Assets

Summary measure: Value of non-land household assets: An aggregate of:

1. **Durable assets**: Sum of the respondent's estimate of the value of household holdings of each asset of different types, if they were to sell them today in their current condition[16] and the respondent's estimate of the value of any shares household members

[16]Note that this variable will sum over the values of two classes of assets in our survey. First, we asked all households to report quantities and values of 'core' assets that they own. Second, we asked households to report the value of a randomly selected asset from a list of 8 'randomised' assets, though households reported quantities for all of these assets. For the 'randomised' assets without a value, we impute the value based on the reported quantity and the mean value of the asset for the sample (from observations for which we have both variables). The set of 'randomised' assets was chosen based on which assets appear least frequently and had low variance in values in the baseline survey.

have in non-financial assets owned by a group (e.g. chairs or tents owned by a group to rent) (varname: s10q_val_*, s12_assets_1_v-s12_assets_54_v).[17]

2. **Livestock:** Total value of the respondent's estimate of how much money they would get if they sold all mature and immature livestock of different types that they own today. We ask about each livestock type (and mature and immature livestock) individually, and then sum these variables (varname: s14r6a_mat*, s14r6a_immat*, s14r6c1_mat_wc*, s14r6c1_immat_wc*, s14r6c1_bees_wc*, s14r6c_mat*, s14r6c_immat*).
3. **Savings:** Total value of savings of all household members held at home, with friends and neighbours, with shopkeepers, with microcredit groups, in mobile money accounts, and in bank accounts. This includes total value of all household ROSCA shares (varname: s8q1* s10g12, s10g13, s10g15).[18]
4. **Net financial liabilities:** Total value of loans taken minus loans given by the household. Loans taken includes those taken from banks, microcredit organizations, other forms of nonprofit or savings groups, employers, merchants, family and friends. Loans made includes all loans the household made to others. Total projected interest is included in the loan value. We include loans in the form of food and exclude loans with face value under KES50 (varname: s8q3*, s8q9*).[19]
5. **Stocks of dried maize*:** Value of stocks of dried maize currently owned by the household, valued using unit prices of dried maize obtained from the village price survey (varname: s5r7s_stock_qty, s5r7n_stock_unit) .

Secondary analysis:

1. We examine if any changes in asset stock are driven by consumer durables or productive assets, which are likely to be used in agriculture or enterprise.
   (a) **Non-productive physical non-land assets:** Total value of the the following assets owned by the household: electrical goods (radio, TV, video player), jewellery, cooking pots/pans, solar panels, metal sheets, furniture (cupboards, beds, chairs, tables, mattresses, mats, etc)(varname: s12_assets_v_1 - s12_assets_v_24).

[17]Assets held in a group does not include shares in ROSCAs or merry-go-rounds. The question is asked for all types of community groups household members are involved in.

[18]We ask how many community groups household members takes part in. We then loop over each group and ask what is the purpose of the group. For groups identified as 'ROSCA/VSLA/merry-go-round/SACCO/table-banking/saving group', we ask what type it is – whether it gives out loans (type 1), only rotates the pot (type 2) or it does both (type 3). In order to avoid counting money that the household has already drawn from the pot as savings in a ROSCA, we ask them to report the value of the household's total 'share' (i.e. the amount they expect to receive when there is a payout in this cycle). Note that for type 2 and type 3 ROSCAs, for households that have already received their share, we code ROSCA savings as 0.

[19]The baseline covariate excludes loans in the form of food and includes loans under KES50.

(b) **Productive physical non-land assets:** An aggregate of:

i. **Enterprise assets:** Total value of the following assets owned by the household: sewing machines, cash tills, tools, vehicles, fishing equipment (varname: s12_assets_v_25 - s12_assets_v_34, s12_assets_v_54).[20]

ii. **Agricultural tools:** Total value of the following assets owned by the household: donkey, oxen, cart, boreholes, irrigation equipment, wheelbarrow, machetes, plough, trough, poultry houses (varname: s12_assets_v_35 - s12_assets_v_53).

iii. **Livestock**, defined as above.

2. We do not currently plan on including the value of land holdings in assets. Sales and purchases of land are infrequent, so it is very difficult for most respondents to estimate land value. We will estimate treatment effects on sales of land, purchases of land, and number of plots owned by the household. If we find evidence of treatment effects on any of these variables, we will 1) incorporate an estimate of the value of land holdings into the aggregated asset variable, using hypothetical land valuations and 2) run robustness tests by analysing assets excluding households who sell or purchase land.

Robustness checks:

1. We have objective measures of the quantity of seven durable assets that were relatively easy for the enumerator to verify.[21] We will compare the treatment effects estimated using self-reported quantities for these seven assets with those using quantities observed by the enumerator.
2. We will examine treatment effects where total assets are defined as an index and livestock holdings are defined as Total Tropical Livestock Units (TLU) currently owned by the household.[22] This avoids measurement error in estimating the value of livestock.

7.3 Revenue

Summary measure - Total household revenue: An aggregate of the following::

[20] We also ask about value of enterprise assets in the enterprise module. This is defined as the aggregate replacement value of all inventory and fixed assets used by any non-agricultural enterprises owned or operated by household members (varname: s6g15a*). We will use this second measure to test the robustness of our estimates.

[21] Only for surveys conducted at the home of the respondent, we have objective measures of the quantities of the following assets in the household: cooking pots and pans, jerry cans, chairs/sofa, tables, radio, TV, poultry house.

[22] We ask the respondent about ownership of mature and immature livestock of different types, and then creating a total TLU measure with this information. One TLU is equivalent to 250 kg of live weight and is used to convert quantities of different animals into a single comparable unit.

1. **Revenue from agriculture:** An aggregate of the following:
   (a) **Household crop production:** Total value of production for each crop the household grew in each of the two rainy seasons, including both production sold and production kept and consumed in-kind.[23] We measure units produced by the household. To value the units produced, the calculation uses the following hierarchy of valuation approaches, relying on the next best approach whenever the required data for the preferred approach is not available for a given crop: (a) the respondent's direct assessment of crop value; (b) the unit price from sales by the same household of the same crop-unit; (c) the crop-unit price obtained from the market price survey; (d) the sub-location median of other households' direct assessments of the value of the same crop-unit; (e) the sub-location median of other households' direct assessments of the value of the same crop, converted using a universal unit conversion ratio;[24] (f) the sub-location median of other households' sales prices of the same crop-unit; (g) the sub-location median of other households' sales prices of the same crop, converted using the universal unit conversion ratio (varname: *crop_code* s5r7a*, s5r7b* s5r7c* s5r7g* s5r7h* s5r7i* village_id_new_combined).[25]
   (b) **Rental income from renting out land across economic activities:** Rent the household received from renting out land or any buildings on it in the last 12 months (varname: s13g2_rplots_rent*).[26]
2. **Revenue from livestock rearing and produce:** An aggregate of the following:
   (a) **Livestock sales:** Total value of animals the household sold in the last 12 months (varname: s14r6k* s14r6l1_wc_* s14r6l_poultry_wc s14r6k_other s14r6l1_other_wc*).[27]
   (b) **Livestock produce:** Total value of livestock production by the household in the last 30 days, including livestock produce that is kept and consumed in-kind, and consists of: beef (cattle meat), sheep/goat meat, poultry meat, other meat,

[23] In the baseline, production data was collected in aggregated form across both seasons. Where the harvest had not concluded in Siaya, respondents were asked to provide projections of the quantity of each crop they expected to harvest.

[24] We will use conversion ratios taken from the market price survey and the field team to convert between local and standard units.

[25] At baseline we do not have hypothetical valuations, so valuations are based on unit prices of sales or on prices from the market price survey (i.e., approaches (a) (d) and (e) are not available).

[26] At baseline, questions are not strictly comparable: i) respondents were asked separately about renting out of plots and renting out of the main compound and ii) for the main compound, respondents were asked about the rental of any buildings on it, but not about rental of land from the main compound. The baseline outcome variable is defined as the sum of rent received for plots and rent received for rental of any buildings on the main compound, both in the last 12 months (varname: s13r4g* s13g2i).

[27] We unfortunately omitted the value of livestock slaughtered for own consumption at endline.

milk, blood, honey, eggs, wool, hides, skin and manure (varname: s14r8a* s14r8e* s14r8f* s14r8g*). We use the reported quantity produced and the price per unit of the livestock goods produced to estimate the value of production in three ways:

i. For livestock products sold by the household, we multiply the quantity sold by the most common price per unit of these sales, which the household reported.
ii. For livestock products that are not sold by the household, we multiply the reported quantity with the most common price the household estimated it would receive if it sold its products.[28]
iii. If household price data is not available, we will value production using local prices, using first prices from the price surveys and then median prices in the sub-location.

3. **Revenue from non-agricultural activities:** An aggregate of the following:
   (a) **Enterprise sales:** Total sales, including value of in-kind income, from any non-farming, non-livestock enterprises owned or operated by household members in the last 30 days the business was in operation (varname: s6g15a_wc, s6r10d).[29]
   (b) **Revenue from community group business activities:*** Total earnings received by any household member from any community group business activity, such as renting out of group assets, in the last 12 months. Group assets include items such as plastic chairs and marquees which could be hired out for events (varname: s10q_inc_*).
   (c) **Income from renting out assets:** Earnings received from renting out any assets owned by the household in the last 12 months, including the value of in-kind payments (varname: s12r1e).
4. **Total household labour earnings:**[30] An aggregate of the following:
   (a) **Household earnings - casual work:** Total monetary value of earnings in cash and in-kind for all household members aged 16 or above from farming another household's land, tending animals owned by other households and in any other casual work in the last 4 weeks (varname: s7r30b_wc).
   (b) **Household earnings - salaried employment:** Total monetary value of earn-

[28] At baseline, we only ask for the unit price of each product if it was sold. To value production in cases where no sale was made, we use prices from the price survey, or if not available, the median unit prices in the sub-location/location.

[29] At baseline, the data excludes sales made on credit and revenue from renting out enterprise assets.

[30] If we see effects on earnings, we will also analyse as secondary outcome the effect on the average daily wage rate. This is total labour earnings from all work outside the household divided by total labour supply outside the household in person days in the last 4 weeks (sum of s7r30b_wc and s7r33b_wc divided by the sum of s7r30a_wc and s7r33a_wc).

ings in cash and in-kind for all household members aged 16 or above from salaried employment working for someone outside the household in the last 4 weeks (varname: s7r33b_wc).

(c) **Remittances received from household members who are migrants*:** Total monetary value of cash and in-kind transfers received in the last 30 days from household members while they were away for work (varname: m14_recieve_amnt_short).[31]

Secondary analysis: We test if changes in production also result in changes in the amount of output sold to generate income. **Summary measure - Income from sales of produce:** An aggregate of the following:

1. **Agricultural sales:** Revenue (in cash or in kind) from the sale of crops in either the long rains season or the short rains season. We define crops as in Section 7.3. For each crop, we ask for the total value of all sales of the crop in each season. (varname: s5r7i*).[32]
2. **Livestock and livestock product sales:** Revenue (in cash or in kind) from the sale of livestock or the sale of livestock products in the last 30 days. We define livestock sales as in Section 7.3. We define livestock products as in Section 7.3. For each livestock product, we ask for the quantity sold in the last 30 days and the unit price of the sale. We then estimate the total revenue received from all product sales in the last 30 days. (varname: s14r8a*, s14r8e*, s14r8f*, s14r8g*, s14r6k*, s14r6l1_wc_*, s14r6l_poultry_wc, s14r6k_other, s14r6l1_other_wc*).
3. **Enterprise sales:** We define enterprise sales as in as in Section 7.3.

Robustness checks: We examine treatment effects when livestock sales are defined in terms of Tropical Livestock Units (TLU), as defined in Section 7.2.

7.4 Investment into economic activity

Summary measure: Total expenditure on investment into economic activity: An aggregate of the following:

1. **Agricultural input expenditures:** An aggregate of the following:
   (a) **Expenditure on the purchase or rental of inputs:** Total expenditure in the last two seasons on fertiliser, seeds and seedlings, insecticide, fungicide, bags and

[31] We only include here what the household members sent back to the household while away. The cost of the household member migration to the household is included in Section 7.4.

[32] We do not include households who have not completed the harvest in Siaya as we do not capture the portion of crops they project they will sell.

storage, farm implements (e.g., ploughs, machetes, hoes), irrigation and pumping equipment, farm machinery and fuel, and fees or interest for farm-related financial services (varname: s5r9b*).

(b) **Rental expenditure on land:** Total rent paid for plots rented in by the household in the last 12 months (varname: s13r5g*).

(c) **Expenditure on hired labour for agriculture:** Total number of days of labour hired in for agriculture in the most recent long rains season multiplied by wages from the village price surveys (varname: s7r13a_long).

2. **Livestock input expenditures:** An aggregate of the following:

(a) **Livestock purchase expenditure:** Total expenditure on purchasing livestock in the last 12 months (varname: s14r6e_wc* s14r6e_poultry_wc s14r6e*_other_wc_*).

(b) **Livestock input expenditure:** Total expenditure on all inputs used in the last 12 months. This includes animal feed, veterinary services, medicines and vaccines, equipment, transportation and construction of livestock enclosures (varname: s14r7b_*).

(c) **Expenditure on hired labour for livestock:** Total number of days of labour hired in for livestock in the last 4 weeks multiplied by wages from the village price surveys (varname: s7r21a_wc)

3. **Non-agricultural expenditures:** An aggregate of the following:

(a) **Enterprise expenditures:** Total expenditures by any non-farming, non-livestock enterprises owned or operated by household members in the last 30 days. Expenditures excludes durable goods/assets but includes maintenance (of land, building, equipment or other fixed assets of the business), buying stocks and inventory, paying wages, fuel and rent costs (varname: s6g17*).[33]

(b) **Spending on community group economic activities*:** Total contribution by the household to all group assets, initiatives or any other economic activities over the last 3 months. This excludes the ROSCA or savings pot (varname: s10g6_v3_*).

(c) **Household members' migration expenditure*:** Total cost of trips made by any household member for work in the last 30 days which required spending more than 2 nights away from home, plus any money or value of in kind support provided by the household while the household member was away (varname:

[33]At baseline, data was collected for individual cost categories separately. The baseline aggregate thus includes the following cost items not explicitly prompted for in the endline aggregate: interest and insurance expenditures; security expenditures; telephone and internet expenditures; public taxes, fees, and fines.

m10_cost_short, m12_send_amnt_short).

Robustness checks:

1. We examine the total area of land under cultivation, measured as the total land in acres under agricultural cultivation by the household, instead of rental expenditure on land. This is an alternative way of capturing changes in use of land as an input. Acreage is aggregated across the short and the long rains seasons (i.e. the same land cultivated in both seasons is counted twice). The baseline variable collected the total area of land allocated to each crop across both growing seasons. We sum up the areas allocated to each crop to obtain the baseline analogue of total land under cultivation across both seasons (varname: area_long, area_short).
2. We examine treatment effects when livestock purchases are defined in terms of Tropical Livestock Units (TLU).

7.5 Human capital investment

All analysis uses the household-level average value for all age-eligible members. We will carry out analysis for members aged 6-20 as age-eligible and also for those aged 14-20 as age-eligible as enrollment for younger children may be subject to ceiling effects. Households with no age-eligible members are omitted from the analysis.

Summary measure - Education index: An index of the following two variables:

1. **Education expenditure:** Total expenditure on education (including school and activity fees, other school related supplies and uniform cost) in the January to December 2017 school year and since the start of the 2018 school year for all children, divided by total number of children in the household (varname: s4g5o_wc_1_201718, s4g5o_wc_2_201718, s4g5p_wc_1_201718, s4g5p_wc_2_201718).[34]
2. **Participation:** Total number of school days attended over the last five days school was in session, divided by the number of children in the household. This is set to zero for non-enrolled children and missing for children living in other cities or enrolled in boarding schools (varname: s4g5k).

Secondary analysis:

1. We estimate all treatment effects above by child gender.

[34] At baseline, education expenditure was asked only for the current academic year.

2. We examine effects on the **Dietary diversity score:** An integer score (maximum 10) indicating the number of food groups from which there has been non-zero consumption of core food items within the household during the 7 day recall period. The 10 food groups are Cereals, Other Starchy Foods, Vegetables, Meat and Poultry, Eggs, Fish and Other Seafood, Beans, Milk, Cooking Fat, and Sugar (varname: s2b*; see Appendix B, Table 2 for list of items included).[35]

7.6 Labour supply

Summary measure - Total household labour supply:[36] An index of the following:

1. **Agriculture in last long rains:**[37] Total days worked by household members in the most recent long rains season (s7r15a_long).
2. **Livestock:** Total days worked by household members in the last 4 weeks for all activities related to livestock raised or owned by the household (varname: s7r25a).
3. **Non-farm enterprises:** Total days worked by household members in the last 4 weeks in the household enterprise(s) (varname: s7r23a).
4. **Labour outside the household:** An aggregate of the following:
   (a) **Casual work:** Total days worked by household members in the last 4 weeks in farming another household's land, tending animals owned by other households and in any other casual work (varname: s7r30a_wc).[38]
   (b) **Salaried employment:** Total days worked by household members in the last 4 weeks in any salaried employment for someone outside the household (varname: s7r33a_wc).
   (c) **Migration:** Total days worked away from the household by all household members who spent at least 2 nights away in the last 30 days for work (varname: m3_days).

Secondary analysis: We examine respondent-level labour supply, as opposed to household-level labour supply. Treatment effects on these outcomes may differ if there is differential measurement error on respondents reports of other household members labour supply or if

[35] Egg consumption was not measured at baseline, so the baseline DDS excludes eggs and is measured on the scale 0-9.

[36] For all activities other than work on household enterprise and on agriculture, days worked by household members was asked for the last 12 months at baseline. This will be scaled to make it comparable to the endline.

[37] We asked for combined labour supply for the short and long rains season at baseline.

[38] Note that at baseline we asked for each activity separately while at endline we asked about all activities together.

different household members respond in different ways to the treatments.

7.7 Consumption

Consumption expenditure (including own-valuation of items consumed from own-production, gifts and transfers) was measured through a short consumption module following Beegle et al. (2012) and Pape and Mistiaen (2015). All respondents were asked about core food and non-durable non-food items; non-core items were allocated randomly across respondents to allow for imputation of values and thus scaling of the aggregates.[39] A comprehensive list of durable items was included,[40] as was a list of social expenditures.

Summary measure - Total household consumption expenditure: An aggregate of household food consumption, expenditure on non-durable and durable household goods and transport and social expenditures and per-child schooling expenditure, all scaled to 30 days, made up of:[41]

1. **Food consumption:** Value of household consumption of 18 core food items (scaled by imputed factor[42] to account for consumption of non-core items) and outside-household food consumption (7 day recall, scaled to 30 days) (varname: s2r5b3*, s2r4f, s2r4g; see Appendix B, Table 3 for list of core food items and Table 4 for list of non-core items included).
2. **Non-food non-durable consumption:** Value of household consumption of nine core non-food non-durable items including household goods, fuel, hairdressing and transport fares, scaled by imputed factor to account for consumption of non-core items (30 day

[39]The core items were chosen to maximise consumption shares captured, based on Kenya Integrated Household Budget Survey (KIHBS) 2005/06 consumption data in the study regions as well as baseline and piloting of this study. Each respondent was randomly allocated and asked about one of 29 groups of three non-core food items and one of 31 non-core non-food items. To construct the food and non-food non-durable aggregates, we are likely to use simple scale factors, following Beegle et al. (2012). For each extra food item, we 1) compute consumption value as the proportion of core food consumption for all households where that item was asked 2) calculate the mean of these proportions, over all households where that item was asked 3) sum the mean proportions over all extra food items (which we interpret as the typical value of extra items consumed as a proportion of the value of core food consumption) and 4) scale food consumption accordingly to estimate total food consumption. The same approach is used for the extra non-food non-durable items. However, we will compare the statistical properties of this approach with the statistical properties of alternative imputation methods, for example multiple imputation.

[40]The list was based on the durable item list in the 2005/06 Kenya Integrated Household Budget Survey, shortened by condensing items into clear categories where possible.

[41]We do not include housing expenditure (on rent) in the aggregate. Consumption measures usually include flow spending e.g. on rent (Deaton and Zaidi, 2002), but here rental markets are very thin so it is difficult to value living in one's own home.

[42]Three of 86 non-core food items were allocated to each household at random to allow imputation of the scale factor.

recall) (varname: s1r2a*; see Appendix B, Table 5 for list of core non-food, non-durable items and Table 6 for list of non-core items included).

3. **Expenditure on durable goods:** Value of household expenditure on durable items and the maintenance of durable items (12 month recall, scaled to 30 days) (varname: s1r3a*; see Appendix B, Table 7 for list of items included).
4. **Social expenditure:** Value of household expenditure on charitable donations, worship contributions, social and entertainment expenditures, weddings and bride price. Regular worship contributions are 30 days recall, wedding expenditures including bride price are since-intervention recall[43] and the other items are all 12 months recall. All are scaled to 30 days (varname: s1r3a*, s1r2a*, s1r2e, s1r2k; see Appendix B, Tables 5 and 7 for list of items included).
5. **Education expenditure:**[44] Expenditure on education (including school and activity fees, other school related supplies and uniform cost) in the January to December 2017 school year and since the start of the 2018 school year for all children. We scale this to 30 days, adjusting for seasonality in education expenditures. (varname: s4g5o_wc_1_201718, s4g5o_wc_2_201718, s4g5p_wc_1_201718, s4g5p_wc_2_201718).[45]

Secondary analysis: We use the standard consumption aggregate excluding housing, defined above, as our summary measure to test for the overall effects of each intervention on consumption. However, the standard consumption aggregate does not map exactly to consumption in our conceptual framework, where we define 'consumption' as expenditure that generates mainly utility returns and 'investment' as expenditure whose primary purpose is to generate future financial returns. Of course, some expenditure may generate both utility and financial returns e.g. expenditure on building may include building a shed for storage of goods to sell, as well as repairing one's house, while consumer durables may be used in businesses.

We thus conduct secondary analysis where we create two aggregates which map more closely onto 'consumption' and 'consumption which may also generate financial returns' (these should be considered alongside our aggregate of investment into economic activity, which captures 'investment'). The conceptual framework suggests these two aggregates may respond

[43] These were 12 months recall at baseline.

[44] This is included in the consumption aggregate so it is comparable with other studies, but the individual variable will be analysed separately in Section 7.5 and we will not include it when doing corrections for multiple testing for variables in the consumption aggregate.

[45] At baseline, education expenditure was asked only for the current academic year. We scale this to 30 days by dividing by the number of months between the start of the school year and the baseline.

to treatment in different ways. We report impact on each variable separately, adjusting for multiple testing within each aggregate.

1. **Expenditure with potential investment returns:** An aggregate of:
   (a) **Expenditure on durable goods**
   (b) **Housing investment expenditure:** Household expenditure on repair and maintenance of house and construction of house or room since intervention (varname: s13f_a, s13f_b).[46] These are usually excluded from standard consumption aggregates (Deaton and Zaidi, 2002).
2. **Expenditure with mainly utility returns:** An aggregate of:[47]
   (a) **Food consumption**
   (b) **Non-food non-durable consumption**
   (c) **Social expenditure**

Robustness checks: We examine an index of housing quality as a fieldworker-verified robustness check for the housing expenditure measure. We collected data on the type of materials used to construct the roof (7 materials), walls (10 materials) and floor (5 materials) of the house, which were observed by the enumerator. The field team then ranked these in order of quality, from worst to best, and gave an explanation for their ranking. We used these to categorise materials into groups, based on whether they are 'improved' from the lowest quality materials. This gave two quality category scores for roof and floor materials (0 or 1), and three for wall materials (0, 1, 2).[48]

8 Definitions of variables to capture mechanisms

We test potential mechanisms through which the intervention(s) may have an effect. We adapt psychological scales and tasks widely used in developed country contexts for the Western Kenyan context. We transform all psychological scales to Z-scores (using the endline mean and standard deviation of the placebo group) to enable comparison across the

[46] At baseline, expenditure on closely related items were measured with 12-month recall: repair and maintenance of the house; decorations, improvements and additions to the house.

[47] Effects on the composition of spending will reflect "Engel effects"– the income elasticities of various consumption items.

[48] Roof materials are scored as follows: 0 = leaves, grass or tins; 1=iron, cement/concrete, tiles or asbestos. Walls material scores are as follows: 0 = mud, mud and poles, or unburnt bricks with mud; 1 = iron/tin sheets, wood, or mud and cement; 2 = unburnt bricks and cement, burnt/stabilised brick, cement/concrete blocks, or concrete and stones. Floor materials are scored as follows: 0 = mud/earth, other organic, or part organic, part finished; 1 = wood, cement or tiles. Our method is based on the method outlined in Arias and De Vos (1996). We then sum these variables to give an housing quality index from 0 to 4, to be interpreted as the number of qualitative improvements made to the house.

outcomes.

8.1 Expectations for one's future outcomes

We measure expectations and aspirations for the future in different domains of life. We use the dimensions – income, asset wealth and children's education – from Bernard and Taffesse (2014) but leave out social status. For all expectations and aspirations questions, assets are defined as "the worth of your house, your furniture, consumer goods like a TV and fridge and any transport vehicles". Monthly income is defined as "all sources of cash income for your household, including what you earn from all agricultural and non-agricultural activities, and money that you have received from any NGO or government programmes." Before we ask about expectations and aspirations, we remind the respondent of the current value of their assets based on their previous entries to anchor their answers.[49] We ask first about aspirations and then about expectations.

Summary measure – Expectations for one's future outcomes: An index of:

1. Expectations for a randomly selected child's attained years of education from the question "What level of education do you think *child name* will achieve?"(varname: s4g5d).[50]
2. Estimate for future assets from the question "What is the level of assets that you think your household will reach at the end of the next ten years?" (varname: asp38a_assets_estimate).
3. Estimate for future monthly income from the question "What is the level of monthly income that you think your household will reach at the end of the next ten years? (varname: asp24a_inc_estimate_month).
4. Mean expected assets at the end of the next ten years, calculated using a log normal or other distribution, taking into account the expected minimum, expected maximum, and the expected probability of earning between the minimum and $k = 1$, between $k = 1$ and $k = 2$, between $k = 2$ and $k = 3$, and between $k = 3$ and $k = 4$ (also the maximum) (this measure is discussed below) (varname: asp39a_assets_min, asp40a_assets_max, asp44_assets1_chance, asp45_assets2_chance, asp46_assets3_chance,

[49] We change the way in which we calculate the respondent's current assets for the endline. At baseline, we obtain this figure by adding the value of the respondent's land, property and any other assets, while at endline we ask the respondent directly for a total.

[50] This question is similar to the question in the fourth round of the Indonesian Family Life Survey. At endline, we select one child (between the age of 6 and 17) in every household closest to the age of 14. At baseline, we asked about all children in the household between the ages of 6 and 24, so we select the child asked about at endline.

asp47_assets4_chance).

5. Mean expected monthly income at the end of the next ten years, calculated using a log normal or other distribution as for assets (varname: asp25a_inc_min_month, asp26a_inc_max_month, asp30_inc1_chance, asp31_inc2_chance, asp32_inc3_chance, asp33_inc4_chance)

To calculate 4 and 5, we elicit probabilistic subjective expectations, as in Dominitz and Manski (1997) and McKenzie, Gibson, and Stillman (2013) (among others), for income and assets at the end of the next ten year period. We ask respondents about how likely they think it is that the realisations of the variable fall within certain intervals. We fit subjective probability distributions for each respondent for each variable. Specifically, for each of income and assets:

1. The enumerator asks respondents for their estimate of the minimum and maximum value of each variable (e.g. "What is the MINIMUM level of annual income that you think your household will reach within ten years? Please suggest a value, where you think that every amount below this amount is impossible.")[51]
2. Based on the minimum and maximum values entered by the enumerator, surveyCTO calculates three respondent-specific thresholds that split the range into four intervals of equal length (e.g. if minimum is 10 and maximum is 30, then the thresholds are 15, 20, and 25).
3. The enumerator draws the values of each threshold on a visual aid. They give each respondent 10 buttons and ask them to allocate to the four intervals, with the number of buttons representing how likely the respondent thinks it is that the realisation of the variable will fall into a given interval. The respondent is required to use all 10 buttons.
4. From these answers, we calculate (separately for income and for assets) $F_{i,k} = P(y_i < Y_{i,k}|\phi_i)$, the probability respondent i assigns to the event that the realisation of the variable falls below threshold $Y_{i,k}$, given her current information ϕ_i. The threshold levels are $k = 1, 2, 3, 4$ and correspond to the three intermediate interval thresholds and the maximum value. To get from intervals to F_{ik}, the points on the CDF, we assume that at Y_4, the threshold that corresponds to respondents' estimate of the

[51] Evidence on whether respondents actually give the minimum or maximum is mixed: Delavande, Giné, and McKenzie (2011a) find that the majority of respondents indeed assign zero probability to realisations beyond these bounds, and that the probability mass that they do allocate to these out-of-bound intervals never exceeds 10 percent. In contrast, Dominitz and Manski (1997) find that respondents treat the minimum as the 20th percentile and the maximum as the 80th percentile, and Delavande, Giné, and McKenzie (2011b) suggest that people report their 90th or 95th percentile as the maximum. However, neither paper points out when asking about the minimum and maximum that these values mean that there is zero probability for all realisations below and above, respectively. In our questionnaire, we explicitly point that out.

maximum value, the CDF is 1. We impose that $P(y_i < Y_{i,min}|\phi_i) = 0$.

To fit distributions, we will likely use the approach of Dominitz and Manski (1997) in the variation used by McKenzie, Gibson, and Stillman (2013), where they fit lognormal distributions by solving the non-linear least-square problem:

$$\min_{\mu,\sigma^2} \sum_{k=1}^{4} \Big(F_{i,k} - G(Y_{i,k}; \mu, \sigma^2) \Big)^2, \tag{4}$$

where $G(\cdot)$ is the CDF of the log-normal distribution and $log(Y_i) \sim N(\mu, \sigma^2)$. However, modelling beliefs is complex, theory seldom gives precise guidance on distributions, and we will explore alternative distributions, partly based on tests of over-identifying restrictions.

Secondary analysis:

1. We specify an indicator equal to one if the respondent expects their child to complete an undergraduate degree or above.
2. We examine the standard deviation for expected income (assets) at the end of the next ten years, calculated using a log normal (or other) distribution and the quantities above, to assess if respondents' uncertainty about their future outcomes changed.
3. We examine changes in other percentiles of the distribution of subjective expectations. For example, even if respondents do not change their mean expectations, they may change their expectation that higher outcomes are possible.

The psychological intervention might also affect expectations about the state of the world, in particular the returns to investment.[52] However, we stress that these were very difficult to measure with illiterate respondents and may be very noisy, making it difficult to detect effects.

Summary measure – Expectations about future returns: An index of the following:

1. **Increase in yields from use of fertiliser**: Respondents' estimate of the percentage increase in production from use of fertiliser. Respondents are asked the amount of dry maize they would harvest in the next long rains season from a one acre plot like most other plots in their area, not using fertiliser, and the amount they would harvest

[52] It is possible that these expectations change in the cash intervention group, for example if there are credit constraints and people were not able to make high-return but lumpy investments, but the main purpose of this test is to understand the mechanism behind any psychological effects.

when applying 50 kg of DAP (the most commonly used fertiliser locally) per acre while planting. They are told the DAP is free and it is the best, official, quality of DAP (varnames: asp79_usual_estimate, asp89_urea_estimate).

2. **Increase in yields from more agricultural labour**:* Respondents' estimate of the percentage increase in production from 12 hours of extra labour on their farm per week. Respondents are asked the amount of dry maize they would harvest in the next long rains season from a one acre plot like most other plots in their area, not using fertiliser, and the amount they would earn if working 12 hours more per week.[53] (varnames: asp79_usual_estimate, asp_90_urea_5hours, asp_91_urea_10hours).
3. **Increase in income from investment in education**:* Respondents' estimate of the percentage increase in monthly income (at age 30) that their child closest to 14 (as in the aspirations question) would see if they finished a university degree, compared to leaving school at the end of secondary schooling (form 4) with a KCSE certificate. (varnames: s4g5_inc_uni, s4g5_inc_4).

8.2 Aspirations for one's future outcomes

Summary measure – Aspirations for one's future outcomes: An index of:

1. Aspirations for a randomly selected child's years of education attained from the question 'What level of schooling would you like *child name* to achieve?" The child is the same as in the expectations question (varname: s4g5c).
2. Aspirations for future assets from the question "What is the level of assets that you would like your household to reach at the end of the next 10 years?" (varname: asp36_assets_desired).
3. Aspirations for future income from the question "What is the level of monthly income that you would like your household to reach at the end of the next 10 years?" (varname: asp23_inc_desired10_month).‡

Secondary analysis:

1. We create an index which can be compared between midline and endline using equation 2, as only the following questions were asked at midline and endline:
   (a) Aspirations for years of education attained by the oldest child of the respondent aged 18 or less from the question "What level of schooling would you like *child*

[53] We also ask the expected increase in harvest and earnings if they worked 18 hours more per week, allowing us to test for nonlinear returns to labour inputs.

name to achieve?" (varname: s4g5c_manipkid)

(b) Aspirations for future income.

2. For both primary and secondary analysis, we specify an indicator variable equal to one if the respondent's desired level of education for their child is an undergraduate degree or above.

8.3 Self-beliefs

Summary measure: Index of self-beliefs: We create an index of the following three scales:[54]

1. The Schwarzer and Jerusalem (1995) scale of **self-efficacy**:‡ Respondents rate the extent to which each statement is true for them from 1 (not at all true) to 4 (completely true). The final composite score at midline and endline is the sum of seven items, ranging from 7 to 28, with high scores indicating high general self-efficacy.[55] (varname: se_*).
2. an adapted version of the 6-item Implicit Theories of Intelligence scale (Blackwell, Trzesniewski, and Dweck, 2007) to measure **growth mindset**:*,‡.[56] Respondents are asked to state the extent to which they disagree with each statement on a scale of 1 (agree strongly) to 6 (disagree strongly). The final score is the sum of the 6 responses for each scale (from 6-36). There are 3 fixed mindset and 3 growth mindset items. We reverse code the 3 growth mindset items such that the higher the score, the more the respondent has a growth mindset (varname: fixed_*, growth_*).
3. The Internal sub-scale from the Internal, Powerful Others and Chance (IPC) scale (Levenson, 1981) to measure **internal locus of control**:[57] The respondents are asked to state the extent to which they agree with each statement on a scale from 1 (disagree strongly) to 4 (agree strongly). The final score is the sum of the 5 responses for each

[54]Clustering these scales together is theoretically justified. Some psychologists argue that self-efficacy and locus of control are similar aspects of the same underlying concept, called "core self-evaluation" (Martocchio and Judge, 1997; Judge et al., 2002b,a). Growth mindset is also argued to be related to self-efficacy, although it is distinct from it, in that having a growth mindset enables people's sense of self-efficacy to survive setbacks (Dweck and Master, 2009). Similarly, growth mindset is related to locus of control, in that both are linked to more perceived control over outcomes and events (Dweck and Leggett, 1988).

[55]Due to time constraints, we reduced the original 10-item scale following baseline data collection, dropping the 3 items which were most highly correlated with others in the scale. At baseline, we use the sum of all 10 items.

[56]We replace intelligence with skillset as the focus of the measure to make the scale applicable in our context of a rural, adult population.

[57]Due to time constraints, we only include 5 out of the 8 statements. We chose measures most relevant to the Kenyan context (excluding, for example, items about driving a car).

scale (from 5-20) (varname: loc_11, loc_1, loc_4, loc_8, loc_14).

Secondary analysis:

1. We analyse the Chance sub-scale of the IPC scale (varname: loc_15, loc_5, loc_9, loc_3, loc_10).
2. We create an index of growth mindset and self-efficacy, as a proxy for the self-beliefs index, which can be compared between midline and endline using equation 2, as locus of control was not asked at midline.

8.4 Time preferences

[58] **Time preferences*:**. We use a standard Multiple Price List (MPL) to measure time preferences over money (Coller and Williams, 1999).[59] We ask respondents to make a choice 7 times between two amounts offered early or later. We do this twice: in the near time frame, we offer money tomorrow or in 15 days and in the future time-frame, money in 15 days versus in 29 days. The amount offered at the earlier date is always equal to KSh 400, while the amount offered at the later date increases from KSh 360 to 1600.[60] We create an index of the following:

1. **Present bias:*** Indicator for if respondent switches to the (higher) future amount later in the near time frame (tomorrow vs. 15 days), than in the future time frame (15 vs. 29 days) (varname: time1 - time7).
2. **Discount factor[61]:*** We assume a linear utility function in money and measure the discount factor using the switch from receiving money soon to later[62]. For example, we ask respondents if they would prefer to receive 400 in 15 days or 440 in 29 days. If

[58]For both time and risk preferences, since poverty is correlated with high discount rates and high levels of risk aversion (Dohmen et al., 2011; Haushofer and Fehr, 2014; Pender, 1996), we will code the variables defined above such that higher values correspond with 'better' outcomes – not present biased, a high discount factor and low risk aversion.

[59]We do incentivised tasks with the respondents to elicit time and risk preferences. Respondents are told that the computer will draw a lottery and there is 1 in 20 chance that they could actually be paid for these tasks. They are told that there is an equal chance for either the time question or risk question to be randomly selected by the computer for payment. Among the time questions, they are told there is an equal chance for each choice to be randomly selected for payment. Half the time the time preference task appears first and the other half the risk elicitation task appears first. We will control for the order in which the questions appear. Respondents are informed whether they have won anything, the amount won and when they should expect to receive the amount at the end of the survey. All payments are made via M-Pesa.

[60]We randomize whether the respondents make the decisions in the near or future time frame first.

[61]Following Andersen et al. (2008) we use the future time frame for the discount factor to account for any transaction costs or additional risk of future income.

[62]In case a respondent switches multiple times, we will use the first switch point

they choose 440, then they have a two-week discount factor between 0.91 and 1 and we assign them the mid-point of 0.96. If they chose 400, then we check their choice between 400 in 15 days and 700 in 29 days and so on until they switch (varname: time1 - time7).

8.5 Risk preferences

Risk preferences*: We use the method developed by Eckel and Grossman (2002). Respondents are asked to make one choice from 6 gambles that are presented to them. Each gamble has two choices A and B. There is 50% chance of receiving a low payoff (A) and 50% chance of receiving a high payoff (B). The amount the respondent will get for the option they choose will depend on whether A or B is randomly chosen by the computer. One gamble gives a certain return and the other choices increase linearly in risk (as measured by the standard deviation).[63] Under the assumption of constant relative risk aversion (CRRA), the gamble chosen by the respondent corresponds to a coefficient of relative risk aversion and wealth level. This allows for the identification of varying levels of risk aversion. We use the rank of the choices from 1-6, increasing in level of risk (varname: risk1).[64]

8.6 Depression

Depression severity score: the 10-item Centre for Epidemiological Studies Depression Scale (CES-D) scale (Andresen et al., 1994). Respondents are asked to give the frequency with which they experience each of the items from 1 (rarely or none of the time, or up to 1 days a week) to 4 (all of the time, or 5-7 days a week). The total score is a sum of all 10 items, with scoring on questions 5 and 8 (positive affect items) reversed. We rebase the score to a scale score ranging from 0 to 30. We then reverse-code the scale, such that higher scores reflect lower depressive symptoms and better mood, to aid comparison with other tests of mechanisms.

Secondary analysis: We examine robustness to a different scaling of the measure, a dummy for whether or not the respondent is depressed. We generate a binary variable where individuals with scores at or above a threshold are identified as at high risk of depression or as experiencing psychological distress. This is how the score is used if it is used for clinical

[63] We randomized the order of the certain and other choices.

[64] We examine robustness to calculating the coefficient of partial risk aversion and report estimates if they differ. Binswanger (1982) reported little impact on analysis of using this as compared to the simple ranking.

screening.[65]

8.7 Cognitive bandwidth

Summary measure - Cognitive bandwidth: An index made up of the following:[66]

1. **Working memory (digit span)*:** Where no response is correct, scored as zero. Otherwise, scored as the length of the longest sequence that respondents can correctly recall, minus two points (varname: dst_*). This is an element from the Wechsler Adult Intelligence Scale (Wechsler, 1958) to measure working memory.
2. **Fluid intelligence (Raven's matrices)* (Raven, 1990):** Scored as the number of correct responses to six patterns provided within the time limit of 30 seconds for completing each pattern (varname: rav_*).
3. **Cognitive control (Numerical Stroop)* (Stroop, 1992; Mani et al., 2013):** Scored as the number of correct responses provided to 3 tasks of 25 number sequences each within the time limit of 30 seconds for each task (varname: stro_c_*, stro_i_*).

We only test cognitive outcomes in the midline survey, just before the placebo or psychological intervention is administered. Because outcomes are measured pre-treatment, no group has yet received their psychological intervention. We test for the effect of being assigned to receive the cash transfer on this index. The control group is the pooled psychological intervention and the placebo group. The treatment group is those in either the cash group, pooled together.

8.8 Tests of alternative mechanisms – level of information

We examine one index of two alternative tests of potential mechanisms through which the psychological intervention might have effects. We do not expect differences between groups in these tests because we sought to ensure there were no differences in the information

[65]Thresholds are not absolute: they are generated when a depression scale is validated by examining the extent to which the scale predicts a diagnosis of depression made in a full clinical assessment in the particular sample, minimising errors of both inclusion and exclusion. In the original CES-D 10 validation, a cutoff of 8 to 10 performed best (Andresen et al., 1994). However, optimal cutoffs varied considerably in other US and Chinese studies, from 8 to 16 (Irwin, Artin, and Oxman, 1999; Boey, 1999; Björgvinsson et al., 2013; Cheng and Chan, 2005). In the only study in sub-Saharan Africa we have found which validates the measure against a full clinical assessment (Baron, Davies, and Lund, 2017), 13 is used. Other studies in sub-Saharan Africa use a threshold of 10 without comparing it to full clinical assessments (Peltzer, Pengpid, and Tiembre, 2013; Pengpid, Peltzer, and Skaal, 2013; Asante and Andoh-Arthur, 2015; Kilburn et al., 2016). We have not conducted such a validation, so examine scores against thresholds of 10 and 13. We will alter thresholds if more studies from the context become available to inform this choice.

[66]These are defined as in https://www.socialscienceregistry.org/docs/analysisplan/1473/document.

provided by the psychological and placebo interventions. First, the psychological intervention may simply encourage mimicry of the actitivites in the video. Second, the psychological intervention may improve recall of information which is contained in both videos. Even if the interventions give the same information, potentially the psychological intervention is more engaging.

Summary measure - Information index: An index of the following:

1. **Mimicry of videos**: An index comprised of dummy variables coded to one if the respondent engaged in any of the following activities at endline, all of which are featured in the videos: (a) weaved baskets; (b) kept savings in a jar; (c) attended a sewing class; (d) trained as a teacher; (e) grew vegetables to sell on the market. (varname: emulation; components: s5s_acts*). We examine whether the video works simply by giving people *specific* information about new activities.
2. **Information recall**: Straight after intervention, at the midline, we test if information recall of specific information contained in both videos, about the returns to education for Kenyan men, is different between the two groups.[67] Both videos contain identical, specific information about the returns to education (varname: end30_*).

Secondary analysis: We also conduct descriptive analysis on two questions, whether there were any names mentioned in the movie and if yes, the name of one of the main characters in the movie. This provides a check on whether the psychological interventions were memorable enough that respondents still remembered the characters (varname: recall_name1_yn, recall_judy).

8.9 Testing additional behavioural mechanisms

We conduct exploratory analysis to test whether the psychological intervention affects decisions in income-generating activities, even if households cannot afford new investment in economic activities or investments have not yet yielded returns. We examine whether households are seeking more information or investing in new technologies or production techniques.

Summary measure - Technology adoption and information seeking*: Sum of the following dummy variables, coded to one if respondent has engaged in the activity since the video intervention:[68]

[67] See https://www.socialscienceregistry.org/docs/analysisplan/1211/document.

[68] We will simply add up the individual items. If we see treatment effects, we will examine the four variables

1. **Technology adoption in crop agriculture*:** (a) saved water for irrigation or used a pump; (b) improved soil e.g. using compost, fertiliser, lime, or ash; (c) used insecticide, fungicide, herbicide, or pesticide; (d) used improved or high yield seeds or seedlings; (e) terraced land or install a fence; (f) tried out a new crop; (g) used lime or ash for the soil, or had the soil tested; (h) engaged in intercropping, row planting, or crop rotation; (i) used a new crop storage solution or greenhouse; (j) used new machine or new farming tool. (varname: s5s_acts*).
2. **Technology adoption in livestock related activities*:** Sum of the following dummy variables, coded to one if respondent has engaged in the activity since the video intervention: (a) vaccinated livestock, used a vet, purchased livestock medicine, or inseminated livestock; (b) sprayed or dipped livestock; (c) built or used a livestock enclosure; (d) tried out any new animal not previously tended to. (varname: s5s_acts*).
3. **Enterprise expansion activities*:** Sum of the following dummy variables, coded to one if, during the last 12 months, a non-agricultural enterprise owned or operated by a household member: (a) introduced new or significantly improved products or services; (b) went into a new market or accessed new customers. (varname: ent_innov*). Note this value may be more than two if there are multiple enterprises in a household.
4. **Information seeking activities*:** Sum of the following dummy variables, each of which is coded to one if, since the video intervention: (a) the respondent asked advice from an extension agent or other farmer; (b) a household member attended at least one training by the One Acre Fund; (c) a phone number associated with household enrolled in an agricultural advisory service offered by Safaricom. (varname: oaf*, s5s_acts*).

8.10 Subsidiary analysis

8.10.1 Transfers

We estimate treatment effects on financial in- and outflows. These are not a core part of our conceptual framework, so we may report these results in this paper or in companion work focused on spillover effects of these interventions.

1. **Transfers given*:** Sum of (a) the total value of any money, food or goods of value less than KES 4,000 given to any person outside the household in the last 30 days and (b) the total value of any money, food or goods of value over KES 4,000 given to any

– technology for crop agriculture, livestock and enterprise, and information seeking – separately to see which is driving effects, and correct for multiple testing over these four variables.

person outside the household since the intervention (varname: s1r3_c_*, s1r3_a_*).

2. **Transfers received:** Sum of the following three sub-components: (a) the total value of any money, food or goods of value less than KES 4,000 received from any person outside the household in the last 30 days; (b) the total value of any money, food or goods of value over KES 4,000 received from any person outside the household since the intervention; (c) the total value of any pension, insurance, scholarship, government grant or lottery received by the household in the last 30 days. (varname: s1r3_d_*, s1r3_b_*, s9_b).[69]

8.10.2 Household composition

We estimate treatment effects on the following variables capturing household composition. This analysis informs the interpretation of effects on outcomes like education, labour and consumption.

1. Net change in number of household members since baseline: Number of household members currently in household minus number of household members at baseline.
2. Number of household members aged 0 to 4, by gender.
3. Number of household members aged 5 to 14, by gender.
4. Number of household members aged 16 and above, by gender.
5. Number of household members aged 6 to 20, by gender.
6. Number of household members aged 6 to 13, by gender.
7. Number of household members aged 14 to 20, by gender.
8. For those who are no longer members of the household, we will analyze the reason for moving.

Households may have split since the baseline survey. In such cases, we attempt to separately survey both the primary male and female respondent households. In cases where we were able to survey both, we will give a sampling weight of 0.5 to each of the two offshoots of the original household and 1 to all other households. If we do not survey both households, we will give the household we find a weight of 1.

[69] The baseline analogue is the sum of: (a) The total value of any money or goods received from persons who are not members of the household in the last 12 months and (b) the total value of any pension, insurance, scholarship, government grant or lottery received by the household in the last 12 months.

9 Extensions

9.1 Extension Analysis

Given the scope of the experiment and data collection, we anticipate that some extension analysis will be included in companion papers rather than the main paper.

First, the discussion in Appendix A abstracts away from heterogeneity in the delivery of the psychological intervention. Some respondents participated in the psychological intervention before receiving their first lump sum cash transfer and some participated after. In addition, both the psychological and placebo interventions were administered in small groups for roughly half the sample. Participants were randomly assigned to groups within villages conditional on some covariates. Companion work will explore if outcomes vary with the relative timing of the two interventions and explore outcomes across respondents assigned to groups with different compositions.

Second, the sketch model treats prices and interest rates as exogenous. Our primary specifications will adjust for aggregate inflation from baseline to endline but will not adjust for spatial price variation within a survey round. We have also collected paper-based market price data, which has not yet been captured. Companion work will estimate between-village price effects using this data (see https://www.socialscienceregistry.org/trials/1484 for details). If we find large price effects, we will proceed as discussed in Section 7.1.

Third, households may interact through markets (e.g. employment, trade) and networks (e.g. informal risk-sharing, altruistic transfers, informal taxation). We do not plan to adjust for market- or network-based spillover effects in this paper. Companion work will explore spillover effects using data on networks and outcomes for respondents in our sample and an additional sample of treatment-ineligible households. If we find substantial spillover effects, we will adjust our interpretation of treatment effects and/or condition on the channels through which spillovers occur.

Fourth, we plan to collect further data on the education of children living in the recipients' households. This will cover test scores, school participation, time use, aspirations, expectations, and potentially measures of intergenerational bargaining over education investments. We will register a new pre-analysis plan for these data before analyzing them.

9.2 Exploratory Analysis

This section describes several analyses that we will explore but are not pre-committing to include in the current paper.

9.2.1 Compliance and compliance-adjusted treatment effects

We will explore compliance and compliance-adjusted treatment effects as a secondary analysis. For each of the four *assigned* treatment groups, we will report the share of respondents *receiving* each of the following six treatment combinations: no cash, psychological or placebo intervention; no cash or psychological intervention but the placebo intervention; no cash or placebo intervention but the psychological intervention; cash but no psychological or placebo intervention; cash and the placebo intervention but no psychological intervention; and cash and the psychological intervention but no placebo intervention. (By design of the experiment, no respondent may receive both the placebo and psychological interventions.)

We will estimate treatment-on-the-treated estimates of the cash transfer and psychological intervention for the full population using the model

$$\begin{aligned} Y_{iv} =& \text{CashReceived}_v \cdot \beta_C^{IV} + \text{PsychReceived}_v \cdot \beta_P^{IV} + \\ & \text{CashReceived}_v \cdot \text{PsychReceived}_v \cdot \beta_{CP}^{IV} + \mathbf{X_{iv}} \cdot \mathbf{\Gamma^{IV}} + \epsilon_{iv}, \end{aligned} \tag{5}$$

instrumenting the treatment receipt indicators (CashReceived, PsychReceived, CashReceived · PsychReceived) with the treatment assignment indicators (Cash, Psych, Cash · Psych). This analysis does not distinguish between respondents who receive the placebo intervention and receive neither the placebo nor the psychological intervention. The experimental design cannot identify the treatment effect of the placebo intervention relative to no intervention without additional assumptions.

In response to a funder request, we will also analyse the marginal effect of adding the psychological intervention *in the population of cash transfer compliers* using the model

$$Y_{iv} = \text{Psych}_v \cdot \beta_P^{CC} + \mathbf{X_{iv}} \cdot \mathbf{\Gamma^{CC}} + \epsilon_{iv}. \tag{6}$$

We will estimate this model on households who were among the first batch of transfer recipients in their village. This omits households that were not assigned to receive cash transfers, households that refused cash transfers, and households that initially refused transfers but

changed their mind when the psychological intervention was already underway. β_P^{CC} in this model measures the average effect of assignment to psychological interventions conditional on accepting a cash transfer. This assesses whether psychological interventions will enhance the effects of cash transfers in a particularly policy-relevant population: compliers with cash transfers. This approach takes compliance as given and does not seek to model the compliance process or compare the population of compliers to the broader population.

9.2.2 Other exploratory analysis

- We will examine whether we can estimate the rate of return to investment and treatment effects on this rate. We will approach this by aggregating all money-metric investments and returns, regressing returns on investments instrumented by the cash treatment, and testing if this rate of return differs across respondents assigned to the psychological and placebo groups. Our basic conceptual framework predicts that this rate of return should be higher for respondents in the psychological treatment group, though this prediction can be flipped by other model features like incorrectly optimistic beliefs. This approach relies on constructing reasonably accurate and comprehensive money-metric aggregate investment and returns, which will be difficult if, for example, a large share of the investment is in children's human capital.
- Our endline survey asks cash transfer recipients how they spent the transfer, which household member received the transfer, and whether they believed that the transfer came with conditions. We will report descriptive statistics for these data and may use them to understand patterns of treatment effects between people receiving both cash and the active psychological treatment and people receiving cash and the placebo treatment. Questions examine the spending breakdown across different categories, in whose name they got the cash transfer, if the respondent believes there were conditions attached and what conditions were, and if the respondent thinks there were consequences of breaking conditions.
- For respondents who completed the psychological or placebo intervention in a group, we collect several measures of their subsequent engagement with group members. We will explore how much interaction occurs and whether the interactions vary by treatment group to understand whether group dynamics might mediate treatment effects. Questions ask if participants in group psych/placebo treatments remember another group member and/or talk to them.

We will run analysis[70] on the following variables but think it is unlikely we will find treatment effects because the outcome is rare (1-2), only applicable to the small subset of households with age-eligible members (3-6) or data is missing because children are not available (7):

1. **University enrollment:** An indicator variable for a household member (up to the age of 24) enrolled in university (varname: s4g5j).
2. **Respondent enrollment:***,‡ An indicator variable if the primary respondent reports enrolling in any school/college/university for education or in vocational training since the intervention (varname: s4_respb).
3. **Took KCPE:*** An indicator variable if a household member took the Kenyan Certificate of Primary Education test (KCPE) in 2017 (varname: s4_kcpe_otherhh).
4. **Took KCSE:*** An indicator variable if a household member took the Kenyan Certificate of Secondary Education test (KCSE) in 2017 (varname: s4_kcse_otherhh).
5. **KCPE score:*** Total score on Kenyan Certificate of Primary Education test (KCPE) for a household member who took the test in 2017 (varname: s4_kcpee).[71]
6. **KCSE grade:*** Average grade on Kenyan Certificate of Secondary Education test (KCSE) for a household member who took the test in 2017 (varname: s4_kcsee).[71]
7. **Anthropometrics:*** We measure the height of all available children under the age of 5 (varname: u5_height). Each child is measured twice and the enumerator records both readings. We will take the average of the two readings and will convert this to z scores using the *zanthro* program in Stata. This is based on the LMS method developed by Cole (1990) and Cole and Green (1992) and we will use the World Health Organisation Child Growth Standards (WHO, 2006) as reference data.

[70] Households with no age-eligible members are omitted from the analysis.

[71] It is unlikely that any household will have more than one member who has taken the test in 2017 but in case there is, we will use the average score.

A Conceptual framework

We sketch a conceptual framework that suggests hypotheses and outcomes for analysis.[72] Suppose households maximize intertemporal expected utility subject to a resource constraint in each period, as in generic models as described in Deaton (1992). Households start the initial period with an endowment of assets. Let $A_{i,t}$ be household i's total asset endowment at t. Assets can be consumed today or invested in ways that deliver future returns (e.g. investment in capital goods which delivers interest; investment in inputs into production that yield revenue).[73] There are three economic constructs of interest in each period in this framework: *assets*, *consumption* and *investment*.[74] The rate of return to investment will depend on the production functions for activities and interest rates on financial assets. The relationships between assets, consumption, and investment depend on preferences and the rate of return to investment.

This simple framework can be extended in several dimensions. We can introduce a fourth economic construct *labour*, allowing households to allocate time in each period between labour and leisure. Current labour will generate assets in future periods by contributing to household production or earning wages.[75] We can introduce investment into *different types of assets* with asset-specific returns. For example, households may invest in both agricultural production and livestock rearing and keep some financial savings. Households will then divide investment between different types based on the type-specific rates of return, which in turn depend partly on the type-specific production functions. We may use the ideas around uncertainty and different asset types to investigate a portfolio choice model. In this model, our experiment will cause households to shift investments between domains with different

[72]The sketch framework informed which variables we measured, how variables are grouped together, and our strategy for accounting for testing multiple hypotheses. It also suggests some predictions about the sign of effects. However, we do not commit to using a specific theoretical model: the experiment is motivated by a broad class of models and this broad framework could be adapted or extended to help interpret particular findings. Predicted signs and magnitudes of effects may vary depending on the specific model used.

[73]We consider human capital, as education of the household's children, also as an asset in which one may invest, although returns will come much later.

[74]We define 'investment' as expenditure whose primary purpose is to generate a future income stream. Given that we study a rural setting, investment may be in productive activities (for example, agriculture, livestock or non-farm enterprises) as well as financial assets. Expenditure on consumer durables or housing has characteristics of both consumption and investment expenditure, as discussed in Section 7.7, which we account for by conducting some tests where we group this expenditure separately from other consumption.

[75]As this is an agricultural setting, land, labour and other assets are allocated on the farm at the beginning of the season, yielding a return at harvest time, which then can be turned into consumption or invested into more assets. For simplicity, we assume no spot cash wage market that yields returns to labour before there are returns to assets.

risk profiles. We can introduce *uncertainty* about future returns to current investment, such that households that invest or work more in period t may nonetheless have lower assets in future periods than households that invest or work less.

Let r_{t+1} be the aggregate rate of return to assets carried over from period t. When household expectations are correct and there is no uncertainty, this is the *true* rate of return to investment and/or labour supply. There may be a wedge between *perceived* and actual returns. For example, poor information may lead to artificially low or high *perceived* returns, which distorts decisions and lowers households' discounted lifetime utility.

We incorporate an additional parameter γ_i, a household-specific parameter describing the household's beliefs about what returns the particular household can achieve, to the rate of return. The perceived rate of return in our model is thus $\gamma_i * r_{i,t+1}$. γ_i is a reduced-form representation of a variety of mechanisms through which households form perceptions of their opportunities and likely future outcomes. Any of these mechanisms may introduce a wedge between perceived and actual returns.

We outline some potential mechanisms which may determine γ_i, although we do not yet model them formally. In any standard model, people may have incorrect expectations about the likelihood of potential outcomes. We also test some mechanisms which draw on literature in psychology. People may have low self-efficacy – beliefs about their own capabilities (Bandura, 1997) – or a fixed mindset – beliefs that one's capacity and life conditions cannot be altered through effort (Dweck, 2012). They may have very unrealistically low or high aspirations for the future (Genicot and Ray, 2017; Dalton, Ghosal, and Mani, 2016). They may have an external locus of control and believe their outcomes are largely determined by fate rather than effort (Lefcourt, 1991).[76] People may be depressed and overly pessimistic about future outcomes (de Quidt and Haushofer, 2018). They may have low cognitive bandwidth and have limited attention to devote to processing information and thinking through choices, leading to suboptimal decision-making (Mani et al., 2013; Shah, Mullainathan, and Shafir, 2012). Measures for these mechanisms are described in Section 8.

Preferences likely also affect choices about investment, but do not enter the model through γ_i. A lower discount rate alters allocation choices between periods, increasing investment and future asset holdings, and reducing consumption in periods near the intervention (t or $t+1$). Risk aversion could enter the model by introducing uncertainty about future returns

[76]We refer to these four beliefs about oneself as "self-beliefs". While psychologists draw distinctions between them, we make a theoretical argument for grouping them in Section 8.

to current investment or assets with asset-specific returns.

Our basic framework treats the household as a unitary actor and does not analyze intra-household allocations of consumption, labour supply, investment, returns on investment, or assets.[77] We also abstract from impacts on the local economy, such as price effects. See Section 9 for further discussion on this issue.

A.1 Effects of the cash transfer

The cash transfer $T_{i,t}$ is added to $A_{i,t}$ and is a "windfall" to lifetime resources for treated households. There are more resources for current consumption or current investment that can yield future consumption. After the transfer, we expect:

- *Higher investment* and *higher asset holdings*;
- *Higher consumption* in both periods (if consumption is a normal good in each period), due to the income effect of the cash transfer; and
- *Lower labour supply* and more leisure in each period (if leisure is a normal good, and also if there are no labour and other market frictions, which may be unlikely).[78]

Households will divide the transfer between current consumption and investment based on the rate of return to investment and the intertemporal utility function. The sizes of the consumption and investment effects depend on preferences, the nature of returns and any frictions in the economy.[79]

This discussion is in terms of "planned" consumption at the time of the transfer, so outcomes "planned" at t for $t+1$. We expect that, relative to the control group, both consumption and assets holdings will increase immediately after the intervention, at t. If risky events affect treatment and control similarly, increased consumption and asset holdings will also

[77] We look at household decision making process for married respondents only in a companion paper.

[78] Prior research often finds zero or positive effects of cash transfers on labour supply. There are multiple extensions of our basic conceptual framework consistent with this pattern. For example, if there are frictions in the labour market and the credit market, then the optimal response to the windfall may be to supply more household labour to activities, increasing labour supply, despite the income effect, and even if leisure is a normal good. This further abstracts from any production function complementarities between labour supply and investment or utility function complementarities between leisure and consumption.

[79] If, for example, households face low returns to investment, we expect large increases in durable and non-durable consumption relative to increases in productive assets in period $t+1$. If there are credit constraints as well as relatively high investment returns, then we expect a large effect on investment relative to consumption and hence a large increase in assets at time $t+1$. If there are credit constraints for consumption, and a pre-existing expectation of higher incomes in the future, then we would observe high increases in consumption spending or durables holdings at $t+1$ and less change in productive assets. For now, we do not formally incorporate credit constraints in the framework.

be observed in future periods from $t+1$ (at endline). However, if, for example, the cash transfer leads to increased investment in one particular asset class, say livestock, that then suffers disease, we would not see an increase in asset holdings relative to the control group, even though there was increased investment.

Expectations and aspirations for one's potential outcomes in the future are likely to increase after a windfall in income. We do not make strong predictions for the effect of income changes on self-beliefs, as this is a new area of study. Increases in income are likely to reduce depression (Baird, de Hoop, and Ozler, 2013; Haushofer and Shapiro, 2016). Poverty reduces cognitive bandwidth (Mani et al., 2013; Shah et al., 2013); so relieving poverty may increase bandwidth, which may account for some changes in decision-making. Changes in income may change preferences: poverty is correlated with high discount rates and high levels of risk aversion (Dohmen et al., 2011; Haushofer and Fehr, 2014; Pender, 1996).[80] However, cash transfers did not affect time preference in Blattman, Jamison, and Sheridan (2017).

A.2 Effects of the psychological intervention

Here, we outline potential economic effects *if* the psychological intervention alters households' perceptions of their future opportunities, including opportunities to which they can allocate their initial endowment of time and capital. This alters γ_i for treated agents and changes the perceived period $s > t$ return to investments made in period t and household allocations to consumption and investment. However, the psychological intervention may not have this effect – this is a hypothesis to be tested. It may also have effects on expectations, preferences or other psychological characteristics. Both the model and predictions might need to be adapted depending on the effects of the interventions.

The intervention will not necessarily induce $\gamma = 1$. It is possible that γ is initially low (high) and the intervention does not raise (lower) it all the way to one. The intervention may also raise (lower) γ past one. Whenever $\gamma \neq 1$, household decisions on $C_{i,t}$ and $A_{i,t+1}$ are still suboptimal. To the extent that the psychological intervention moves the perceived rate of return closer to (respectively farther from) the actual rate of return, the intervention reduces (respectively increases) distortions and raises (respectively reduces) households' discounted lifetime utility.[81] It may be difficult to detect overall welfare implications at t or $t+1$,

[80] Debates remain whether this is due to ingrained preferences or a reflection of heterogenous constraints or shifts due to changing contexts (e.g. preferences measured over money reflecting background income).

[81] We discuss in Section 9 some extensions that we may explore to differentiate whether our intervention

but these would be captured in differences between treatment and control groups in welfare outcomes, such as in consumption, in longer term follow-ups.

We discuss the effects on economic outcomes in the case where households had a low γ_i and the intervention increases it to be closer to $\gamma = 1$. The shift in γ_i increases the opportunity cost of consuming today.[82] This generates both an *income effect* – the value of current assets increases because they yield higher returns in the future – and a *price effect* – the opportunity cost of current consumption increases. The cash transfer discussed above has only an income effect.

The income effect of raising γ_i is identical to the cash transfer: current consumption and assets rise and labour supplies falls; period $t+1$ assets rise. The substitution effect of raising γ_i lowers current consumption and raises current investment and period $t+1$ assets. The magnitude of the negative effect on consumption and positive effect on asset holdings in each period depends on preferences and market conditions such as the intertemporal substitution elasticity, interest rates, and discount rates.[83]

Thus, if the psychological intervention raises γ_i:

- We would expect *increased investment* at t into productive assets and activities to take advantage of higher future returns, and *higher asset holdings* from t onwards. Higher perceived returns would also make a more risky set of activities attractive for given risk preferences. It is possible, but less likely, that a response to higher perceived returns may lead, with a large income effect, to lower investments and more consumption at t.[84]
- The effect on *consumption* in the short run (whether at t or $t+1$, at endline) is ambiguous, depending on preferences and market conditions, due to the interplay of

affects perceived or actual rates of return or both. However, differentiating these interpretations is not our primary goal.

[82] While conceptually different from how we defined γ_i, a change in the discount rate may theoretically have effects which are similar to a change in γ_i. We test for which mechanism is at work using empirical tests of alternate mechanisms described in Section8.

[83] If there is also a flow of labour income in each period or another source of income not affected by interest rates *directly*, then there will also be a further negative "endowment" or "wealth" effect on consumption and a positive effect on asset holdings in each period. The present value in each period t of future (labour) earnings (i.e. the discounted value) will be lower if γ_i has increased, driving down consumption at t. The same occurs at $t+1$ relative to future periods, pushing consumption, relatively speaking, further to the future (Deaton, 1992).

[84] Predictions on future asset holdings are again in "in expectation," as seen from t. Furthermore, if γ_i increases more than 'true' returns, holdings in the future may become lower than the control group, if investments are poorly allocated.

income and substitution effects. For example, where $\gamma_i * r_{t+1}$ is larger than the discount rate, the substitution and endowment effects together are likely to dominate the income effect and investment will dominate consumption.

- The effect on *labour supply* is ambiguous. We can also interpret γ_i as a multiplicative shifter of the economic return to labour supply. For example, the psychological intervention might increase workers' perceptions of their own productivity, raising their market wage. Higher γ_i generates a substitution effect that increases labour supply relative to leisure and an income effect that increases leisure relative to labour supply.[85]

The effects of an increase in γ_i on psychological outcomes after some time are less clear. We expect the intervention to increase self-beliefs, aspirations and expectations for future outcomes at t. But by $t+1$, these effects may fade out or change. Experiences which are independent of the intervention (for example, failures or successes) may cause people to revise self-beliefs, aspirations and expectations.

We do not predict any effects of the psychological intervention on preferences or depression and have not mapped out predictions for these cases. The intervention is not likely to be intensive enough to affect depression or preferences over the period of study.[86]

Of course, even if our analysis of mechanisms suggests that γ_i changes, we may not see behavioural changes: this may be due to various market failures, not least credit market failures. Households may know there are better returns, but they may not be able to act on this due to credit constraints.

A.3 Effects of combining the interventions

One treatment group are offered both the cash transfer and the psychological intervention. The overall constraint at the start of period $t+1$ is $A_{i,t+1} = (1+\gamma_i * r_{t+1}) * (A_t - C_t)$. There is both a cash transfer $T_{i,t}$ at t (adding resources to be divided between more consumption and carrying resources over to the future) and a psychological intervention $\Delta\gamma_i$ (raising $A_{i,t+1}$ by raising investment returns). In this model, the total effect of the combination intervention is not just the sum of each intervention. There may also be an additional effect from the higher returns to the cash transfer, which adds a further income effect $\Delta\gamma_i * T_{i,t}$. Households

[85] This prediction becomes even more complicated if there are frictions in labour or other markets, or labour and financial investment are complements or substitutes in agricultural production, livestock rearing, etc.

[86] Longer, more intense psychological interventions do increase patience (Alan and Ertac, 2018; Blattman, Jamison, and Sheridan, 2017) but light-touch interventions tend not to alter preferences (Bernard et al., 2014; Haushofer, John, and Orkin, 2018).

receiving both treatments believe cash will have a higher return in future if invested than households in the cash only arm do. We may expect at least as high an impact on *investment and asset holdings at t* in the treated group as the sum of the impacts observed in each of the two treatment arms. However, this is a hypothesis to be tested – it may be that overall effects are not larger than the sum of each of the effects.

This combined additional effect *may* be large in two cases. First, those receiving only the cash transfer may invest poorly as they are not considering profitable opportunities and the psychological intervention opens their minds to these options. Second, those who receive the psychological intervention still face other constraints, such as credit constraints. They may realise there are better opportunities, but this may not lead to more investments if resources cannot be reallocated from, say, essential consumption. Indeed, the combination may then lead to substantially higher outcomes and the specific effect from $\Delta\gamma_i * T_t$ could be high.

However, the combined intervention may plausibly increase investment by less than the sum of the individual effects. For example, the psychological intervention may increase perceived return to long-term investments like education while the cash transfer relaxes credit constraints to making these investments. The interaction effect on physical assets in $t+1$ can be negative if investment in human capital constitutes a large share of investment. It is also possible that the combined intervention leads to lower *consumption* than the sum of the individual effects. This can occur if period $t+1$ consumption of some goods is an increasing but concave function of $t+1$ assets. In both cases, lifetime asset holdings and lifetime consumption possibilities increases by more than the sum of the psychological and cash transfer intervention effects. However, particular consumption or assets groups observed at $t+1$ may still increase by less than the sum of the impacts of each of the interventions, through rational behavioural responses.

We do not have a clear prediction for complementarity between cash and psychological interventions on expectations or preferences, given there are few clear predictions for the effects of cash on psychological outcomes.

The conceptual framework does not generate predictions about the relative magnitudes of the effects of the cash transfer, psychological intervention, and combined effects. This would require more specific parametric assumptions. We will, however, use the relative magnitudes of the cash transfer and psychological intervention effects to understand the quantitative importance of the psychological intervention.

A.4 Mapping variables to the framework

Table 1 shows the predicted effect of the treatments separately and jointly on the main families of outcomes laid out in Section 6. Predictions for the psychological variables are for positive increases in γ_i that bring it closer to 1.

These predictions ignore uncertainty and assume that our endline occurs late enough to observe returns to all investments, except those in education.[87] We do not directly observe r, the *rate of return* on investment for most investment activities. Instead, we observe levels of investment and levels of return on investment for the activities listed in table 1. We return to this issue in Section 9.

We include one class of variables, 'transfers,' in the table but not in the conceptual framework. This covers measures of non-market financial flows into and out of the household, defined in Section 8.10.1. Prior research on spillover effects of cash transfers suggests that cash transfers increase informal financial outflows and have ambiguous effects on financial inflows (for example, higher wealth from a cash transfer might decrease redistributive transfers but increase reciprocal transfers as part of a mutual insurance network.) Neither our framework nor prior empirical work generates predictions about the effect of the psychological intervention or interaction of both interventions on financial in- or outflows.

Table 1 shows the predicted effect of the treatments separately and jointly on variables we broadly categorise as 'psychological'. The model includes both r_{t+1}, the general rate of return to investment in the future across households, and $\gamma_i * r_{t+1}$, the household-specific rate of return. Extensions to the model would allow these returns to differ across economic activities. We measure respondents' *subjective expectations about their future outcomes* on two key dimensions in our model, assets $A_{i,t+k}$ and consumption $C_{i,t+1}$, after ten years (see Section 8.1):[88]

- Physical asset stocks, measured as "the worth of your house, your furniture, consumer goods like a TV and fridge andany transport vehicles".
- Human capital stocks, measured as the level of education a selected child will attain.

[87] These predictions correspond to the 'likely' cases discussed above. But there exist cases of the framework that generate different predictions, not least if there are market imperfections. If, for example, financial investments and labour are technical complements in the agricultural production function and the cash transfer alleviates a credit constraint that kept financial investment in agriculture sub-optimally low, then the cash transfer might increase labour supply. Labour market frictions that make households depend on household labour for economic activities would similarly affect labour supply predictions.

[88] We selected ten years to allow time for agricultural investments to yield returns. Human capital investments in children may not be earning returns by this point.

Table 1: Predicted signs of treatment effects on each class of variables in endline data

Class	Variable	Cash effect	Psych effect*	Interaction effect*
Economic				
Assets	Durable assets	Positive	Positive	Positive
	Livestock			
	Savings			
	Net financial liabilities			
	Dried maize stock			
Consumption	Food consumption	Positive	Ambiguous	Ambiguous
	Non-food non-durable consumption			
	Non-food durable consumption			
	Housing expenditure			
Revenue	Revenue from agriculture	Positive	Positive	Positive
	Revenue from livestock			
	Revenue from Non-Farm Enterprises			
	Revenue from non-household activities**			
Investment activities	Spending on agriculture	Positive	Positive	Positive
	Spending on livestock			
	Spending on NFE inputs & migration			
Labour supply	Agriculture	Decrease	Ambiguous	Ambiguous
	Livestock			
	NFEs			
	Non-household activities*			
Education investment	Spending on education	Positive	Positive	Positive
	Education participation			
Transfers	Received	Ambiguous	Ambiguous	Ambiguous
	Given	Positive	Ambiguous	Ambiguous
Mechanisms				
Expectations	For future outcomes	Positive	Ambiguous	Ambiguous
	For gains from investment			
Aspirations	For future outcomes	Positive	Ambiguous	Ambiguous
Self-beliefs	Self-efficacy	Ambiguous	Ambiguous	Ambiguous
	Locus of control			
	Growth mindset			
Preferences	Discount rate	No effect	No effect	No effect
	Present bias	No effect	No effect	No effect
	Risk aversion	No effect	No effect	No effect
Depression		Positive	No effect	No effect
Cognitive load		Positive	Can't test	Can't test

*These predictions are only for the case where the psychological intervention increases γ_i and moves it closer to 1.

**Includes casual work, salaried work and migration for work.

- Potential cash consumption, measured as expected annual cash income.[89]

We also measure respondents' *expectations about the likely changes in income possible from financial and labour investments in agricultural production*, which can be viewed as proxies for γ_i for specific types of investments.[90]

We also measure aspirations, people's long-term goals for the future, measured as the level of physical and human capital assets and cash consumption they would *like* to achieve in ten years time (as in Bernard and Taffesse (2014)).

We measure "self-beliefs" – self-efficacy, internal locus of control and growth mindset. Self-beliefs do not enter the model directly as they do not map directly on to γ_i. However, γ_i is a function of self-beliefs, so changes in self-beliefs are likely to affect it. We measure self-beliefs directly using psychological scales described in Section 8. We also measure depression using a standard scale and cognitive load and preferences using standard task-based measures.

A.5 Timing in the framework and in our data

We use a discrete framework where agents make consumption, investment, labour supply, and leisure choices in period t and receive investment returns and labour market earnings in period $t+1$. In the rural setting we study, the time lag between investment/labour and returns varies across activities. Labour supply to the market may generate immediate returns. Agricultural production and some livestock production is seasonal, so returns will be realized with a seasonal lag. This timing may also impose a seasonal structure in returns to non-farm enterprises. Returns to investment in human capital and some livestock will be realized much more slowly.

We conduct a baseline survey, then offer cash transfers and/or psychological interventions and then conduct an endline survey approximately one year later. The period between the interventions and endline includes at least one complete agricultural season, from land preparation to harvest. We ask questions about current consumption, assets, and labour, as well as retrospective questions about investments made and returns realized through the agricultural season.

[89] In piloting, this was much easier to estimate than expenditure. It obviously is not a perfect proxy for consumption, as it does not count income from subsistence production, while it does not account for dissavings or savings.

[90] We view results here with caution as these were very difficult to measure with illiterate respondents and may be very noisy.

Given this timing, it is likely that the period covered in our endline data collection will loosely correspond to period t investment and period $t+1$ consumption, labour, investment returns, and assets. However, some long-term investments may not have yielded returns by the time of our endline, creating the appearance of a low rate of return. For example, investments in young children's human capital will definitely not yield monetary returns within a year. We study treatment effects on different types of investments to try and explore this issue. Negative shocks may also result in low returns despite high investments.

B Consumption item lists

Table 2: Dietary Diversity Score Groups and Items

DDS group	DDS group description	Endline code	Item description
1	Cereals	3301	Maize Grain (include grain ground at mill)
		3302	Maize Flour
		3303	Rice
		3305	Bread/Mandazi (only if purchased)
2	Other Starchy Foods	3307	Sweet Potato
		3309	Cooking Banana
3	Vegetables	3312	Tomatoes
		3313	Kale - Sukuma Wiki
		3314	African nightshade (Osuga/Managu)
4	Meat and Poultry	3315	Red Meat (beef/mutton/goat)
		3316	Chicken and other poultry
5	Eggs[91]	3326	Eggs
6	Fish and Other Seafood	3317	Omena
		3318	Fresh Fish
		3319	Dried/Smoked Fish
7	Pulses and Nuts(=Beans)	3310	Beans
8	Milk	3320	Fresh Milk
9	Oils and Fats(=Cooking Fat)	3321	Cooking Fat
10	Sugar	3322	Sugar

Table 3: Core Food Items and Codes (7 day recall)

Baseline code	Endline code	Item description	DDS group
3301	3301	Maize Grain (include grain ground at mill)	1
3302	3302	Maize Flour	1
3303	3303	Rice	1
3305	3305	Bread/Mandazi (only if purchased)	1

[91] Eggs were not included at baseline.

3307	3307	Sweet Potato	2
3309	3309	Cooking Banana	2
3310	3310	Beans	7
3312	3312	Tomatoes	3
3313	3313	Kale - Sukuma Wiki	3
3314	3314	African nightshade (Osuga/Managu)	3
3315	3315	Red Meat (beef/mutton/goat)	4
3316	3316	Chicken and other poultry	4
3317	3317	Omena	6
3318	3318	Fresh Fish	6
3319	3319	Dried/Smoked Fish	6
3320	3320	Fresh Milk	8
3321	3321	Cooking Fat	9
3322	3322	Sugar	10

Table 4: Extra Food Items and Codes (7 day recall)

Baseline code	**Endline code**	**Item description**	**Randomisation group**
-	3326	Eggs	*All*
-	3401	Mushrooms and Asian Vegetables	1
-	3402	Peaches/plums/grapes/strawberries	1
-	3403	Wheat flour	1
-	3404	Tinned or packeted soups	2
-	3405	Cabbages	2
-	3406	Cucumber	2
-	3407	Avocado	3
-	3408	French beans/Runner beans/Broad beans	3
-	3409	Onion (bulbs or leeks)	3
-	3410	Tinned vegetables	4
-	3411	Tomato sauce/chilli sauce	4
-	3412	Yeast, baking powder, bicarbonate (magadi)	4
-	3413	Tea (leaves or bags)	5
-	3414	Melons	5
-	3415	Camel or Rabbit meat	5

-	3416	Jaggery (Nguru)	6
-	3417	Pasta (Spaghetti, Macaroni, Noodles e.g. Indomie)	6
-	3418	Peas (Garden, snap, snow)	6
-	3419	Mangoes/Pineapples/Passion fruits	7
-	3420	Broccoli/Radish/Baby corn/Sweet corn	7
-	3421	Millet (grain or flour)	7
-	3422	Biscuits	8
-	3423	Aubergines-Egg plant (Biringanya)	8
-	3424	Pork	8
-	3425	Lard (From Butcheries)	9
-	3426	Macadamia or Cashew Nuts	9
-	3427	Crisps	9
-	3428	Capsicums (Pilipili hoho)	10
-	3429	Cheese	10
-	3430	Coconut/guavas/tree tomato	10
-	3431	Beef brawn/pork brawn/bacon	11
-	3432	Potatoes (Irish)	11
-	3433	Sausages/Smokies/Hot dog	11
-	3434	Pumpkin/Squash/Butternut/Courgette/ Marrow	12
-	3435	Sesame seeds/Simsim	12
-	3436	Green Maize	12
-	3437	Cowpeas	13
-	3438	Tinned baby food	13
-	3439	Chick(en) peas	13
-	3440	Ripe Banana	14
-	3441	Coffee (instant or ground)	14
-	3442	Tinned fish	14
-	3443	Salt	15
-	3444	UHT (long life), condensed/tinned or powder milk	15
-	3445	Apples/pears/loquats	15
-	3446	Soya, barley or oats (grain or flour)	16

-	3447	Sugar cane	16
-	3448	Arrow Roots-Nduma	16
-	3449	Chocolate/sweets/chewing gum/other confectionary	17
-	3450	Frozen Fish Filets	17
-	3451	Jam/marmalade/honey	17
-	3452	Prawns /Other sea Foods (not fish) (fresh or frozen)	18
-	3453	Pepper - Pilipili	18
-	3454	Grams (all types)	18
-	3455	Carrots	19
-	3456	Lemons/limes/grapefruit/tangerine (NOT oranges)	19
-	3457	Cakes	19
-	3458	Popcorn	20
-	3459	Food seasonings (eg Royco, Knorr etc) and spices (incl pilau masala, ginger-tangawizi)	20
-	3460	Peanut butter	20
-	3461	Peas	21
-	3462	Offal (liver, kidney etc, matumbo)	21
-	3463	Milk sour (Mala) or yogurt (clotted milk)	21
-	3477	Lettuce, celery	22
-	3465	Mustard/pickles/vinegar	22
-	3466	Tinned beans or pulses	22
-	3467	Cocoa/drinking chocolate/soya drink	23
-	3468	Canned meat/ham/salami	23
-	3469	Commercial soft drinks (sodas, mineral water, squashes, health drink, fruit juice etc)	23
-	3470	Spinach	24
-	3471	Beetroot or turnips	24
-	3472	Yams	24
-	3473	Butter, ghee or margarine	25
-	3474	Porridge flour (mixed/fortified)	25
-	3475	Wheat grain	25

-	3476	Coriander leaves (Dania)	26
3323 (Siaya)	3323	Oranges	26*
3324 (Siaya)	3324	Beer	26*
3325 (Siaya)	3325	Pawpaws	27*
3308	3308	Cassava (flour or root)	27*
3311	3311	Groundnuts	27*
3306	3306/3328†	Chapati	28‡
3304	3304	Sorghum (grain or flour)	28‡
-	3328/3306†	Other alcoholic beverages: Spirits, wine, traditional brew (muratina, buzaa, changaa, cider)	28‡
-	3327	Tobacco: Cigarettes, tobacco (raw), tobacco (processed), cigars, snuff, miraa (khat)	29‡

* Moved from core to extra on 30/05/2018 (all questionnaire versions after 1805290010)

† Accidental code swap during endline; addressed in data cleaning.

‡ Moved from core to extra on 07/06/2018 (all questionnaire versions after 1806051904)

Table 5: Core Nonfood Items (30 day recall)

Baseline code	**Endline code**	**Item description**[92]	**Component of aggregate**
3101	3101	Household soap and detergents	Non-durables
3104	3104	Body lotion, petroleum jelly	Non-durables
3106	3106	Haircuts and hairdressing (men and women)	Non-durables
3107	3107	Regular worship contributions	Social
3110	3110	Milling grains and staple crops into flour	Non-durables
3111	3111	Transport fares (buses, matatu, motorbikes, taxis etc)	Non-durables
3112	3112	Mobile phone airtime and charging	Non-durables
3113	3113	Water (bottled/piped/from tank)	Non-durables
3115	3115	Paraffin, kerosene	Non-durables
3116	3116	Charcoal	Non-durables

[92]Item 3105, 'All medicines (fever/painkillers, anti-malaria, anti-worm, herbal, others)' was also included in the questionnaire but is not included in any of the outcome variables.

Table 6: Extra Nonfood Items (30 day recall)

Baseline code	**Endline code**	**Item description**[93]	**Component of aggregate**
-	3501	Garbage/Refuse/Sewerage collection and services	Non-durables
-	3502	Gas, LPG or Biogas	Non-durables
-	3503	Beauty services other than haircuts and hairdressing	Non-durables
-	3504	Napkins/diapers/pullups for infants	Non-durables
-	3505	Non-electric personal grooming equipment (razors, nail cutters, comb, toothbrush etc)	Non-durables
-	3506	Baby oil and baby powder	Non-durables
-	3507	Hair oil/cream, weaves, wigs or hairpieces	Non-durables
-	3508	Handkerchiefs/Serviettes/Pocket tissues/ Wet wipes	Non-durables
-	3509	Make-up, perfume, after-shave	Non-durables
-	3510	Shampoo, conditioner or deodorant	Non-durables
-	3511	Sanitary towels/tampons/cotton wool/panty liners	Non-durables
-	3512	Toothpaste or mouth wash	Non-durables
-	3513	Personal effects (purse, wallet, bag, sunglasses, umbrella, lighter etc)	Non-durables
-	3514	Baby carriage	Non-durables
-	3515	Household products: insecticide, disinfectant, freshener, floor polish	Non-durables
-	3516	Match box, candles	Non-durables
-	3518	Spectacles, contact lenses, dentures	Non-durables
-	3519	Driving lessons, car rental, parking fees	Non-durables
-	3520	Telephone and internet charges (exclude mobile phone)	Non-durables
-	3521	Natural/artificial flowers, flower pots	Non-durables

[93]Item 3517, 'Medical products and equipment (syringe, elastoplast, thermometer, medical gloves, etc)' was also included in the questionnaire but is not included in any of the outcome variables.

-	3522	Expenses for pets (purchase, food, accessories, veterinary fees)	Non-durables
-	3523	Bank, ATM, mobile money and other financial service charges (not interest payments)	Non-durables
-	3524	Souvenirs, jewellery, clocks and watches	Non-durables
3102	3102	Toilet supplies (toilet soap, toilet paper)	Non-durables
3103	3103	Shoe polish/cream	Non-durables
3109	3109	Domestic workers and services (including laundry)	Non-durables
3114	3114	Batteries	Non-durables
3117	3117	Firewood	Non-durables
3118	3118	Petrol or diesel	Non-durables
3119	3119	Electricity	Non-durables

Table 7: Nonfood Items (12 month recall)

Baseline code	**Endline code**	**Item description**[94]	**Component of aggregate**
3201	3201	Women's clothing and footwear	Durables
3202	3202	Men's clothing and footwear	Durables
3203	3203	Children's clothing and footwear (excluding school uniforms)	Durables
3204	3204	Tailoring services, cloth, sewing supplies	Durables
3206	3206	Books, stationery, postal expenses (excluding textbooks and exercise books for education)[95]	Social
3207	3207	Toys, sports equipment, musical instruments, tapes, video, CDs, DVDs, music	Social
3208	3208	Household cleaning equipment (brooms, brushes, etc.)	Durables
3209	3209	Kitchen equipment (pots, sufuria, pans, frying pans, mwiko, buckets, basins, other utensils)[96]	Durables

[94]Items 3205, 'Medical and health services (modern and traditional: doctor, nurse, hospital fees, etc.)', 3226, 'Fines, certificates, legal services', 3227, 'Insurance (all types: life, health, education, vehicle, property, etc.)' and 3228, 'Taxes (all types: income, land, housing, etc.)' were also included in the questionnaire but are not included in any of the outcome variables.

[95]Baseline wording: 'Books, stationery, postal expenses (excluding textbooks)'

[96]Baseline wording: 'Kitchen equipment (pots, pans, frying pans, mwiko, buckets, basins, other utensils)'

3210	3210	Dishes (crockery, cutlery, glassware, etc.)	Durables
3211	3211	Carpets, rugs, curtains, linens (towels, sheets, blankets, etc.)	Durables
3212	3212	Mosquito nets	Durables
3213	3213	Furniture (bed, mattress, chair, table, sofa, etc.)	Durables
3214	3214	Non-electric household goods (jiko/stove, paraffin lamps, etc.)	Durables
3215	3215	Electrical goods including solar (radio, TV, lights, fan, fridge, mixer, etc.)	Durables
3216	3216	Repair and maintenance of household goods	Durables
3217	*(Land module)*	*Household expenditure on repair and maintenance of house since intervention*[97]	(Housing investment)
3218	*(Land module)*	*Household expenditure on construction of house or room since intervention*[98]	(Housing investment)
3220	3220	Vehicle repair, maintenance, parts and licenses (do not include petrol/diesel)	Durables
3221	3221	Entertainment activities, excursions and holidays	Social
3222		Charity and donations (excluding regular worship contributions)	
	3222	Charity and donations (excluding regular worship contributions) in cash to organisations, not individuals	Social
3224	*(Separate module)*	*Household expenditure on marriages of household members since intervention*[99]	Social
3225	*(Separate module)*	*Household expenditure on bride price since intervention* [100]	Social
3229	3229	Membership fees and contributions to clubs (except ROSCAs/merry-go-round)	Social

[97]Baseline wording (12 month recall, item code 3217) 'Repair and maintenance of the house'.

[98]Baseline wording (12 month recall, item code 3218) 'Decorations, improvements and additions to the house'.

[99]Respondents were directed to include 'cost of ceremony, all costs of setting your son/daughter up in their new home, transport to a new location, purchases for the home', but not to include the bride price. Baseline wording (12 month recall, item code 3224) 'Marriages and other ceremonies'.

[100]Respondents were directed to include 'the value of any money, food and goods paid as bride price'. Baseline wording (12 month recall, item code 3225) 'Dowry or bride price'.

References

Alan, S. and S. Ertac. 2018. "Fostering Patience in the Classroom: Results from a Randomised Educational Intervention." *Journal of Political Economy* 126 (5): 1865 – 1911.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström. 2008. "Eliciting Risk and Time Preferences." *Econometrica* 76 (3): 583–618.

Anderson, M. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.

Andresen, E. M., J. A. Malmgren, W. B. Carter, and D. L. Patrick. 1994. "Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D." *American Journal of Preventive Medicine* 10 (2): 77–84.

Arias, E. and S. De Vos. 1996. "Using Housing Items to Indicate Socioeconomic Status: Latin America." *Social Indicators Research* 38 (1): 53–80.

Aronson, J., C. B. Fried, and C. Good. 2002. "Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence." *Journal of Experimental Social Psychology* 38 (2): 113–125.

Asante, K. O. and J. Andoh-Arthur. 2015. "Prevalence and Determinants of Depressive Symptoms Among University Students in Ghana." *Journal of Affective Disorders* 171: 161–166.

Baird, S., J. de Hoop, and B. Ozler. 2013. "Income Shocks and Adolescent Mental Health." *Journal of Human Resources* 48 (2): 370–403.

Bandura, A. 1997. *Self-Efficacy: The Exercise of Control*. New York: W.H. Freeman.

Baron, E. C., T. Davies, and C. Lund. 2017. "Validation of the 10-Item Centre for Epidemiological Studies Depression Scale (CES-D-10) in Zulu, Xhosa and Afrikaans Populations in South Africa." *BMC Psychiatry* 17 (1): 6.

Beegle, K., J. D. Weerdt, J. Friedman, and J. Gibson. 2012. "Methods of Household Consumption Measurement Through Surveys: Experimental Results from Tanzania." *Journal of Development Economics* 98 (1): 3–18.

Benjamini, Y., A. Krieger, and D. Yekutieli. 2006. "Adaptive Linear Step-up Procedures that Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.

Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse. 2014. "The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia." *Centre for the Study of African Economies Working Paper* WRS/2014-16: 1–48.

Bernard, T. and A. S. Taffesse. 2014. "Aspirations: An Approach to Measurement with Validation Using Ethiopian Data." *Journal of African Economies* 23 (2): 189–224.

Binswanger, H. P. 1982. "Empirical Estimation and Use of Risk Preferences: Discussion." *American Journal of Agricultural Economics* 64 (2): 391–393.

Björgvinsson, T., S. J. Kertz, J. S. Bigda-Peyton, K. L. McCoy, and I. M. Aderka. 2013. "Psychometric Properties of the CES-D-10 in a Psychiatric Sample." *Assessment* 20 (4): 429–436.

Blackwell, L. A., K. H. Trzesniewski, and C. S. Dweck. 2007. "Theories of Intelligence and Achievement Across the Junior High School Transition: A Longitudinal Study and an Intervention." *Child Development* 78 (1): 246–63.

Blattman, C., J. C. Jamison, and M. Sheridan. 2017. "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioural Therapy in Liberia." *American Economic Review* 107 (4): 1165–1206.

Boey, K. W. 1999. "Cross-Validation of a Short Form of the CES-D in Chinese Elderly." *International Journal of Geriatric Psychiatry* 14 (8): 608–617.

Cheng, S.-T. and A. C. Chan. 2005. "The Center for Epidemiologic Studies Depression Scale in Older Chinese: Thresholds for Long and Short Forms." *International Journal of Geriatric Psychiatry* 20 (5): 465–470.

Cole, T. J. 1990. "The LMS Method for Constructing Normalized Growth Standards." *European Journal of Clinical Nutrition* 44 (1): 45–60.

Cole, T. J. and P. J. Green. 1992. "Smoothing Reference Centile Curves: the LMS method and Penalized Likelihood." *Statistics in Medicine* 11 (10): 1305–1319.

Coller, M. and M. B. Williams. 1999. "Eliciting Individual Discount Rates." *Experimental Economics* 2 (2): 107–127.

Dalton, P. S., S. Ghosal, and A. Mani. 2016. "Poverty and Aspirations Failure." *The Economic Journal* 126 (590): 165–188.

de Quidt, J. and J. Haushofer. 2018. "Depression Through the Lens of Economics: A Research Agenda." In *The Economics of Poverty Traps*, edited by C. Barrett, M. Carter, and J.-P. Chavas. Chicago, IL: University of Chicago Press, 127–152.

Deaton, A. 1992. *Understanding Consumption*. Oxford UK: Oxford University Press.

Deaton, A. and S. Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." *World Bank Living Standards Measurement Study Working Paper* LSM135: 1–108.

Delavande, A., X. Giné, and D. McKenzie. 2011a. "Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How Sensitive are Answers to Variations in Elicitation Design?" *Journal of Applied Econometrics* 26 (3): 479–497.

———. 2011b. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics* 94 (2): 151–163.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioural Consequences." *Journal of the European Economic Association* 9 (3): 522–550.

Dominitz, J. and C. F. Manski. 1997. "Using Expectations Data to Study Subjective Income Expectations." *Journal of the American Statistical Association* 92 (439): 855–867.

Duckworth, A. L., T. A. Kirby, A. Gollwitzer, and G. Oettingen. 2013. "From Fantasy to Action Mental Contrasting With Implementation Intentions (MCII) Improves Academic Performance in Children." *Social Psychological and Personality Science* 4 (6): 745–753.

Dweck, C. S. 2012. "Mindsets and Human Nature: Promoting Change in the Middle East, the Schoolyard, the Racial Divide, and Willpower." *American Psychologist* 67 (8): 614–622.

Dweck, C. S. and E. L. Leggett. 1988. "A Social-Cognitive Approach to Motivation and Personality." *Psychological Review* 95 (2): 256–273.

Dweck, C. S. and A. S. Master. 2009. "Self-Theories and Motivation: Student's Beliefs about Intelligence." In *Handbook of Motivation at School*, edited by K. R. Wentzel and D. B. Miele. New York: Routledge, 123–140.

Eckel, C. C. and P. J. Grossman. 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." *Evolution and Human Behaviour* 23 (4): 281–295.

Firpo, S. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75 (1): 259–276.

Genicot, G. and D. Ray. 2017. "Aspirations and Inequality." *Econometrica* 85 (2): 489–519.

Haushofer, J. and E. Fehr. 2014. "On the Psychology of Poverty." *Science* 344 (6186): 862–867.

Haushofer, J., A. John, and K. Orkin. 2018. "What Motivates Health Behavior: Preferences, Constraints, or Beliefs? Evidence from Psychological Interventions in Kenya." *Unpublished* (Princeton University): 1–59.

Haushofer, J. and J. Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *Quarterly Journal of Economics* 131 (4): 1973–2042.

Irwin, M., K. H. Artin, and M. N. Oxman. 1999. "Screening for Depression in the Older Adult: Criterion Validity of the 10-Item Center for Epidemiological Studies Depression Scale (CES-D)." *Archives of Internal Medicine* 159 (15): 1701–1704.

Judge, T. A., J. E. Bono, A. Erez, E. A. Locke, and C. J. Thoresen. 2002a. "The Scientific Merit of Valid Measures of General Concepts: Personality Research and Core Self-Evaluations." In *Psychology of Work: Theoretically Based Empirical Research*, edited by J. M. Brett and F. Drasgow. Mahwah, New Jersey: Lawrence Erlbaum Associates, 55–77.

Judge, T. A., J. E. Bono, A. Erez, and C. J. Thoresen. 2002b. "Are Measures of Self-Esteem, Neuroticism, Locus of Control, and Generalised Self-Efficacy Indicators of a Common Core Construct?" *Journal of Personality and Social Psychology* 83 (3): 693–710.

Kilburn, K., H. Thirumurthy, C. T. Halpern, A. Pettifor, and S. Handa. 2016. "Effects of a Large-Scale Unconditional Cash Transfer Program on Mental Health Outcomes of Young People in Kenya." *Journal of Adolescent Health* 58 (2): 223–229.

King, L. A. 2001. "The Health Benefits of Writing About Life Goals." *Personality and Social Psychology Bulletin* 27 (7): 798–807.

Lee, D. S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76 (3): 1071–1102.

Lefcourt, H. M. 1991. "Locus of Control." In *Measures of Personality and Social Psychological Attitudes*, edited by J. P. Robinson, P. R. Shaver, and L. S. Wrightsman. San Diego: Academic Press, 413–499.

Levenson, H. 1981. "Differentiating Among Internality, Powerful Others, and Chance." In *Research with the Locus of Control Construct*, vol. 1, edited by H. Lefcourt. New York, NY: Academic Press, 15–63.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149): 976–980.

Martocchio, J. J. and T. A. Judge. 1997. "Relationship Between Conscientiousness and Learning in Employee Training: Mediating Influences of Self-Deception and Self-Efficacy." *Journal of Applied Psychology* 82 (5): 764.

McKenzie, D., J. Gibson, and S. Stillman. 2013. "A Land of Milk and Honey with Streets Paved with Gold: Do Emigrants Have Over-Optimistic Expectations About Incomes Abroad?" *Journal of Development Economics* 102: 116 – 127.

Morisano, D., J. B. Hirsh, J. B. Peterson, R. O. Pihl, and B. M. Shore. 2010. "Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance." *Journal of Applied Psychology* 95 (2): 255–264.

Oettingen, G. and P. M. Gollwitzer. 2010. "Strategies of Setting and Implementing Goals." In *Social Psychological Foundations of Clinical Psychology*, edited by J. E. Maddux and J. P. Tangney. New York: The Guilford Press, 114–135.

Oyserman, B. D. . T. K., Daphna. 2006. "Possible Selves and Academic Outcomes: How and when Possible selves Impel Action." *Journal of Personality and Social Psychology* 91 (1): 188–204.

Pape, U. and J. Mistiaen. 2015. "Measuring Household Consumption and Poverty in 60 Minutes: The Mogadishu High Frequency Survey." *World Bank Policy Research Working Paper* 8430: 1–23.

Peltzer, K., S. Pengpid, and I. Tiembre. 2013. "Mental Health, Childhood Abuse and HIV Sexual

Risk Behaviour Among University Students in Ivory Coast." *Annals of General Psychiatry* 12 (18): 1–8.

Pender, J. L. 1996. "Discount Rates and Credit Markets: Theory and Evidence from Rural India." *Journal of Development Economics* 50 (2): 257–296.

Pengpid, S., K. Peltzer, and L. Skaal. 2013. "Mental Health and HIV Sexual Risk Behaviour Among University of Limpopo Students." *South African Journal of Psychiatry* 19 (2): 25–30.

Peters, M. L., I. K. Flink, K. Boersma, and S. J. Linton. 2010. "Manipulating Optimism: Can Imagining a Best Possible Self be Used to Increase Positive Future Expectancies?" *Journal of Positive Psychology* 5 (3): 204–211.

Raven, J. C. 1990. *Standard Progressive Matrices: Sets A, B, C, D & E.* Oxford, UK: Oxford Psychologists Press.

Schwarzer, R. and M. Jerusalem. 1995. *Generalised Self-Efficacy Scale.* Windsor, UK: NFER-NELSON.

Shah, A. K., S. Mullainathan, and E. Shafir. 2012. "Some Consequences of Having Too Little." *Science* 338 (6107): 682–685.

Shah, A. K., S. Mullainathan, E. Shafir, and J. Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 338 (6107): 976–980.

Stadler, G., G. Oettingen, and P. M. Gollwitzer. 2009. "Physical Activity in Women: Effects of a Self-Regulation Intervention." *American Journal of Preventive Medicine* 36 (1): 29–34.

Stroop, J. R. 1992. "Studies of Interference in Serial Verbal Reactions." *Journal of Experimental Psychology: General* 121 (1): 15.

Wechsler, D. 1958. *The Measurement and Appraisal of Adult Intelligence.* Baltimore, US: Williams & Wilkins Co.

WHO. 2006. *WHO Child Growth Standards: Length/Height-For-Age, Weight-For-Age, Weight-For-Length, Weight-For-Height and Body Mass Index-For-Age: Methods and Development.* Geneva: World Health Organization.