Promoting Future Orientation Among Cash Transfer Recipients: Analysis Plan for Manipulation Checks

Kate Orkin*, Robert Garlick†, Johannes Haushofer‡, Richard Sedlmayr§, and Stefan Dercon¶

July 8, 2017

1 Overview

This analysis plan pertains to an initial set of outcomes associated with the trial "Promoting Future Orientation Among Cash Transfer Recipients" conducted in two Kenyan counties, Homa Bay and Siaya.[1] This limited set of outcomes is collected immediately after psychological treatment (or placebo treatment) is administered to trial participants. The objective is to test if the psychological constructs, including economic aspirations, that are presumed to mediate the long-term impacts were in fact successfully manipulated through the psychologically active intervention. We prespecify the sampling protocol, randomization protocol, variable definition and construction, estimation strategy, and hypothesis testing strategy. This analysis plan was lodged before the administration of the psychological and placebo treatments was completed.

A second analysis plan will be lodged before the main endline surveys are completed that will prespecify the variable definition and construction, estimation strategy, and hypothesis

*Department of Economics, Centre for the Study of African Economies and Merton College, University of Oxford. kate.orkin@merton.ox.ac.uk

†Department of Economics, Duke University. robert.garlick@duke.edu

‡Department of Psychology, Department of Economics, and Woodrow Wilson School of Public and International Affairs, Princeton University. haushofer@princeton.edu

§Blavatnik School of Government, University of Oxford. richard.sedlmayr@bsg.ox.ac.uk

¶Blavatnik School of Government, Department of Economics and Centre for the Study of African Economies, University of Oxford. stefan.dercon@economics.ox.ac.uk

[1]See https://www.socialscienceregistry.org/trials/996 for the trial registration.

testing strategy for the outcomes collected in the endline surveys.

2 Interventions

This trial evaluates two interventions individually and in combination. One is an unconditional cash transfer program, the other is a psychological intervention. The selection of psychological intervention, and details of its delivery, were refined through extensive piloting activities.[2] Trial participants who are not assigned to the psychological intervention are instead assigned to a psychologically inactive "placebo" intervention, which has also been piloted. The psychologically active and placebo interventions were delivered separately to each individual in some villages and in randomly selected groups of 2-4 individuals in other villages. (See section 4 for more details.) Those trial participants who are not assigned to the cash condition receive no component of the cash transfer intervention. Additional context on each intervention is provided below.

2.1 Cash Transfer Intervention

The cash transfer intervention is the standard program implemented by GiveDirectly, a nonprofit that operates in Western Kenya. In the course of the registration process, eligible individuals are provided with a mobile money account as well as a mobile phone if they do not yet have these. This is followed by 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 USD450; and a second large transfer ("Lump Sum B") of USD450 minus the price of the mobile phone.

2.2 Psychologically Active Intervention

This intervention integrates three psychological approaches that have been linked to future orientation. Specifically:

1. It promotes a Growth Mindset (Dweck, 2012): the belief that one's capacity and life conditions are malleable through effort, that setbacks and obstacles are an opportunity to learn, and that the process of addressing and overcoming challenges is enjoyable.

[2] See https://www.socialscienceregistry.org/trials/991 for information on these piloting activities.

2. It encourages participants to imagine their Best Possible Selves (King, 2001).
3. It helps participants identify and prioritize goals, and delineate specific strategies for achieving them, following Duckworth et al. (2013), Morisano et al. (2010) and Oettingen and Gollwitzer (2010).

In practice, this intervention involves three core activities:

1. Two back-to-back ten-minute videos, viewed by participants individually on a tablet computer. These portray local women from similar socioeconomic backgrounds who start out with economic difficulties, but respond in future-oriented ways that model the aforemetioned psychological approaches. A brief public service announcement, unrelated to the two videos, is shown between them.
2. A facilitated one hour long exercise that involves a goal drawing exercise and critical discussion. It is meant to help participants translate the approaches modeled in the videos into their own lives.
3. The distribution of a reminder (i.e., a calendar depicting the role models from the movies), as well as a set of stickers that participants may stick on the calendar to represent their goals.

The participants watch the videos and complete the exercises individually in Homa Bay county and in small groups in Siaya county.

2.3 Psychologically Inactive (Placebo) Intervention

The psychologically active intervention may function in part through other channels than the hypothesized ones. For instance, the experience of viewing a video on a tablet or engaging with a facilitator may have an impact even if no valuable content is provided in these interactions. Further, the psychologically active intervention necessarily transfers some content that may alter productivity in ways that are not central to the theory of change (e.g. factual information that may alter people's beliefs, a calendar that may alter people's timeliness, etc.). Therefore, those trial participants who are not assigned to the psychological intervention are instead assigned to a psychologically inactive intervention (in lieu of a pure control group). The videos shown in this intervention use content from the other videos to talk broadly about life in Western Kenya; however, no role models or psychologically active narratives are presented. We include at least one shot of every scene and character from the other videos, including introductory shots of scenery. We exclude any elements that are likely to manipulate critical psychological variables, such as shots of people conveying obvious emotion, eye contact with people, music that creates emotion, or insightful narratives. Placebo participants also participated in an exercise with exactly the same elements: they were reminded of the content of the video, they discussed the facts presented, and they drew the scenes that were most memorable. Placebo participants also receive a calendar. Placebo participants watch the videos and complete the exercises either individually in Homa Bay county and in small groups in Siaya county.

3 Outcomes of Interest

We collect data on five topics in the immediate follow-up survey: self-efficacy, fixed mindset, economic aspirations, education plans and aspirations, and recall of information from the video. Some topics or families include multiple outcome measures and for each family we define one primary outcome, shown in italics in the list below. The name of the variables used in the construction of each measure are in parentheses.

We construct one measure of **self-efficacy**:[3] [4]

1. *scale from 0 to 28, constructed from the sum of 7 items, each scored on a 0 to 4 point scale from Likert questions* (se1_r, se2_r, se3a_r, se6_r, se7_r, se8_r, se9_r)

We construct one measure of **growth mindset**:[5]

1. *scale from 6-36, constructed from the sum of 6 items, each scored on a 1 to 6 point scale from Likert questions* (growth_1, growth_2, growth_3, fixed_1, fixed_2, fixed_3; final three questions are reverse-coded)

We construct two measures of **economic aspirations**, both winsorized at the 99th percentile:

[3]The scale to assess generalised self-efficacy (GSE) was developed via item analysis on pilot data. Items that met any of the following criteria were removed: increased Cronbach's α if item removed, low corrected item-total correlation ($\leq$ 0.25), low loading on primary un-rotated factor (<0.30) and high cross-loading (>0.30), acceptability and comprehension ($\geq$ 20% item non-response) and low item variation ($\geq$ 80% identical responses on the item) (Lamping et al., 2002). As a result of this process, GSE 4,5, 7, 9 and 10 (Schwarzer and Jerusalem, 1995) were removed.

[4]Responses of individual respondents will be dropped if responses on $\geq$ 2 items are missing. When fewer items than this cut off is missing for a particular individual, scores are adjusted to generate homogeneous score ranges using an appropriate multiplier.

[5]Responses of individual respondents will be dropped if responses on $\geq$ 3 items are missing. When fewer items than this cut off is missing for a particular individual, scores are adjusted to generate homogeneous score ranges using an appropriate multiplier.

1. the desired annual household income (asp2_inc_desired)
2. *the desired annual household income minus the previous year's household income* (asp2_inc_desired, asp1_total_inc)

We construct six measures of **education plans and aspirations**:

1. an indicator variable equal to one if the respondent expects their child to enroll in school next year (ed1)
2. an indicator variable equal to one if the respondent's desired level of education for their child is less than completing a university degree (ed2)
3. an indicator variable equal to one if the respondent's desired level of education for their child is at least a complete university undergraduate degree (ed2)
4. an indicator variable equal to one if the respondent's desired level of education for their child is at least a complete professional degree (ed2)
5. an indicator variable equal to one if the respondent expects to enroll in any new form of education or training (own_ed1)
6. *an index constructed by averaging these five measures, using the inverse sample covariance matrix to construct the weights*[6]

We construct three measures of **information recall from the movie**:

1. an indicator variable equal to one if the respondent correctly answers whether any character name was mentioned in the movie[7] (recall_name1_yn)
2. *an indicator variable equal to one if the respondent correctly recalls the name of the lead character from either movie* (recall_judy)
3. an indicator variable equal to one if the respondent correctly recalls the relationship between education and employment for young Kenyan men.[8] (end30_1, end30_2, end30_3, end30_4, end30_5)

We also administer three cognitive tasks prior to the delivery of the intervention: digit span, raven's matrices and numerical stroop. The analysis of these measures will be detailed in a later PAP.

[6]See section 6 for a more detailed explanation of this index construction method.

[7]Names are mentioned in the psychological intervention but not the placebo. So the correct answer is "yes" for treated respondents and "no" for placebo respondents.

[8]This outcome measures how well respondents recall purely factual information from the movie. The active and placebo treatments are designed to contain the same factual information. So we hypothesize that the mean value of this outcome will not differ between the active and placebo treatment groups. If there is a difference, we will interpret this as evidence that factual information and active psychological content are complements in information recall.

4 Sampling and Treatment Assignment

The sampling and treatment assignment scheme is based on household census data collected by the study team in the Homa Bay and Siaya counties in western Kenya. Teams of field officers completed a short census with all consenting households in all villages in target locations. We used these census data to determine village and household eligibility, assign villages to treatment, and draw samples of eligible households in eligible villages. GiveDirectly also conducted a separate census in all villages in Homa Bay and some villages in Siaya assigned to receive cash treatment.

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
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[9]
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 included homeless households and lowered the per capita housing space cutoff to 42,400cm^2 to reflect changes in GiveDirectly's targeting criteria.

We combined some pairs of small villages.[10] We also dropped any village with fewer than

[9] 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.

[10] We aggregated pairs of villages only if they had previously been administered as a single village, were

15 eligible households. Note that GiveDirectly and the research team conducted separate censuses in Homa Bay and part of Siaya and respondents' answers may differ across the 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.

We assigned eligible villages to treatment using a sequential stratified random assignment algorithm. The algorithm was sequential because we assign 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. Location: This is an administrative division in Kenya containing roughly 10-50 villages. We constructed location blocks as pairs of geographically adjacent locations. The first, second, third, and fourth groups of villages contained respectively four, three, three, and two location 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. We then created an indicator variable equal to one if the village amenity index exceeded the sample median.[11]
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 assign villages in each stratification block to the four treatment

geographically contiguous, and each contained fewer than 40 households.

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

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.

In each village, we randomly drew two samples of eligible households: 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 noncompliers or 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.

The active 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. In Siaya, groups of 2-4 people watched the video and completed the exercise together, with a pair of two facilitators.

4.1 Group Assignment

In each village, intervention participants are treated in multiple **sessions**. Participants in every session are assigned to one to three **groups**. As a basic rule, assignment to groups is randomised. Each session therefore corresponds to one execution of the group assignment questionnaire (see Appendix A for group size rules).

Intervention participants are assigned to groups as follows:

1. Prior to the day of the intervention, invitations are made.
   (a) Participants are asked to show up for specific sessions.
   (b) The order in which invitations are initially made is randomised.
2. At the beginning of each session, one of the SFO's uses the group assignment questionnaire a single time to assign all of the participants to groups at random.
   (a) Participants who show up to a session uninvited are allowed to join if we intended to treat them in another session.
   (b) Participants who were invited but don't show up will be invited to another session.
   (c) Group sizes follow pre-specified rules as a function of the number of participants who show up (see the appendix).
   (d) To make group sizes more even, participants might be asked at random to return for a later session (called "rescheduling"). To reduce the inconvenience to participants, rescheduling is not used in the concluding session for the day, unless there are too many who show up. Rescheduling is not used in the very last session held in the village.
3. Participants who show up *late* (after group assignment is completed) can be added to one of the groups at the FO's discretion. This is only allowed before the group questionnaire starts.

5 Estimation

We will estimate models of the form

$$\begin{aligned} Y_{iv} = {} & \text{Cash}_v \cdot \beta_C + \text{Psych}_v \cdot \beta_P + \text{Cash}_v \cdot \text{Psych}_v \cdot \beta_{CP} \\ & + Y_{0iv} \cdot \gamma_0 + T_{iv} \cdot \tau_1 + T_{iv}^2 \cdot \tau_2 + \alpha_v + \epsilon_{iv}, \end{aligned} \tag{1}$$

where i and v index individuals and villages, Y_{iv} denotes the outcome of interest measured in the follow-up, Y_{0iv} denotes the outcome of interest measured in the baseline, $Cash_v$ and $Psych_v$ are indicator variables equal to one for villages assigned to receive respectively cash and psychological treatments, T_{iv} equals the number of days between the follow-up survey and the cash "token" (set to zero if the psychological intervention occurs before the cash transfer), and α_v is a stratification block fixed effect (the the constant term is subsumed into the vector of fixed effects).[12] All outcomes in the immediate follow-up survey are measured for the respondent or one of their children, so i indexes both individuals and households. The parameters of interest are $(\beta_C, \beta_P, \beta_{CP})$, respectively the treatment effects of the cash transfer, the psychological intervention, and both interventions together. We will report bootstrap standard errors for these parameters, using 1000 iterations of a nonparametric bootstrap that resamples villages, stratifying by treatment assignment.

If Y_{0iv} is missing, we will impute the missing value using the sample mean within the stratification block. We will conduct the imputation inside each bootstrap iteration, following Shao and Sitter (1996).

6 Hypothesis Testing

For each outcome listed in section 3, we will test four statistical hypotheses:

1. The cash transfer has no effect on the outcome conditional on receiving a placebo, $\beta_C = 0$.
2. The psychological intervention has no effect on the outcome relative to the placebo, $\beta_P = 0$.
3. The interaction effect of the cash transfer and psychological interventions have the same effect on the outcome as the interaction effect of the cash transfer and placebo interventions, $\beta_{CP} = 0$.
4. The cash transfer and psychological intervention have equal effects on the outcome relative to respectively no cash transfer and the placebo intervention, $\beta_C = \beta_P$.

We expect that both the cash transfers and the psychological intervention will increase self-efficacy, growth mindsets, economic aspirations, and education plans and aspirations $(\beta_C, \beta_P > 0)$. Different models yield different predictions about whether the two interventions are substitutes or complements, so β_{CP} may be positive or negative. We test if β_C is smaller or larger than β_P to compare the cost-effectiveness of the two interventions and provide a benchmark for the size of the psychological intervention effect but this comparison is not an object of primary interest. The active psychological intervention may improve recall of factual information delivered during the video, in which case β_P will be positive for the recall measures.

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

All hypothesis tests will be based on variance-covariance matrices derived using the bootstrap algorithm described in section 5.

We will report both standard p-values and sharpened q-values that control the false discovery rate across outcomes within each of the five variable families from section 3 (Anderson, 2008; Benjamini et al., 2006).[13] [14] Rather than pre-specifying a single q, we report the minimum q-value at which each hypothesis is rejected. For some outcome families we will construct an index following Kling et al. (2007) in three steps. First, we will re-code all primary outcomes so that higher values correspond to "better" outcomes. Second, we will standardize the outcomes to have mean zero and standard deviation one. Third, we will calculate the average of the standardized constituent outcomes, weighted by the inverse covariance matrix. Where an outcome value is missing for a respondent, we will omit this outcome from the index construction. We will then estimate model (1) using this index as the left-hand side variable and test $\beta_C = 0$, $\beta_P = 0$, $\beta_{CP} = 0$, and $\beta_C = \beta_P$.

7 Adjustments for Missing Data in Follow-up Surveys

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). If any one outcome is missing for more than 4% of the sample, we will implement two analyses to characterize the missing data:[15]

1. We will compare the fraction of missing data 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: $(\beta_C, \beta_P, \beta_{CP}, \beta_C - \beta_P, \beta_C - \beta_{CP}, \beta_P - \beta_{CP})$.
2. We will regress a missing data indicator on a vector of baseline covariates using a logit model, report the marginal effects, and test if the marginal effects are individually or jointly significantly different to zero. The baseline covariate vector will be selected using

[13] The self-efficacy and fixed mindset families each contain only one outcome so we will not implement corrections within these families. The information recall family contains three outcomes. We will implement a correction across the first two outcomes but will exclude the outcome about recall of the education-employment relationship; this outcome is measuring a different underlying concept to the other two outcomes and is not expected to move in the same direction.

[14] However, we will not control the false discovery rate across the four hypotheses within each outcome.

[15] Due to a survey version control error, outcome ed3 was not recorded in some villages. We expect more missing data for this outcome.

a logistic LASSO algorithm from respondent age, respondent education at baseline, respondent household income at baseline, the asset index defined in section 4, number of members of the respondents' household, respondent marital status, indicators for the field officers who administered the psychological intervention, sublocation indicators, and indicators for the month of the follow-up survey. We will construct standard errors using 1000 iterations of a nonparametric cluster bootstrap; we will treat the LASSO-selected model as fixed in this bootstrap algorithm.

For any specific outcome where responses are missing for more than 4% of the sample, we will two use analyses to assess the sensitivity of our results to missing data:

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 6. We will construct standard errors using a two-stage bootstrap algorithm where we estimate the both weights and the regression parameters in each bootstrap iteration.
2. We will construct bounds on parameters $(\beta_C, \beta_P, \beta_{CP})$ using the trimming procedure described in Lee (2009).

8 Robustness Checks

We will assess the robustness of the estimated coefficients of model (1) with these additional analyses:

- We will estimate model (1) using weighted least squares to account for missing data, as described in section 7.
- We will estimate Lee bounds on the coefficients $(\beta_C, \beta_P, \beta_{CP})$ from model (1), as described in section 7.
- We will estimate a system of equations using two-stage least squares, where receipt of cash and psychological treatments are instrumented by assignment to respectively cash and psychological treatments. The second stage equation will be

$$\begin{aligned} Y_{iv} = {} & Cash\hat{Received}_{iv} \cdot \beta_C + Psych\hat{Received}_{iv} \cdot \beta_P \\ & + Both\hat{Received}_{iv} \cdot \beta_{CP} + Y_{0iv} \cdot \gamma_0 + T_{iv} \cdot \tau_1 + T_{iv}^2 \cdot \tau_2 + \alpha_v + \epsilon_{iv} \end{aligned} \tag{2}$$

The three first stage equations will use the same right-hand side specification as model

(1) and $PsychReceived_{iv}$, $CashReceived_{iv}$, and $BothReceived_{iv}$ as outcomes.[16] We define a household as receiving a treatment if it receives the treatment at any time between the baseline and endline surveys.

- We will estimate model (1) omitting the covariate vector $(Y_{0iv}, T_{iv}, T_{iv}^2, \alpha_v)$.
- We will estimate model (1) with field officer fixed effects and examine the distribution of field officer fixed effects to see how much the treatment effects vary with across people conducting the intervention. Field officers are not randomly assigned to respondents, so this is not an experimental comparison.
- We will estimate model (1) with linear controls selected using a LASSO algorithm from respondent age, marital status, baseline household income, baseline household assets (using the index described in section 4), baseline education, and baseline household size, field officer fixed effects, month-of-intervention fixed effects, the number of days between receipt of cash and psychological treatments and the interaction of this time difference with an indicator variable equal to one if and only if respondents received the psychological intervention after the cash intervention.[17]
- We will also estimate a system of two equations using the same right-hand side specification as model (1) and respectively a fixed mindset score and growth mindset score as left-hand side variables. The fixed mindset score is the sum of 3 item scores and ranges from 3-18 (variables fixed_1, fixed_2, fixed_3); the growth mindset score is the sum of 3 item scores and ranges from 3-18 (variables growth_1, growth_2, growth_3). This explores whether any treatment effect on the aggregate fixed mindset measure in section 3 is driven by decreases in growth mindset or increases in fixed mindset.

We will assess the robustness of the hypothesis test results (i.e. the p-values described in section 6) with this additional analysis:

- We will construct p-values using a permutation test based on a randomization inference approach. Specifically, we permute the treatment assignments following the algorithm in section 4, estimate model (1) using the permuted treatment assignment, and calculate the test statistics $\hat{t}_C^p$, $\hat{t}_P^p$, and $\hat{t}_{CP}^p$. We will iterate over 1000 permutations and use the empirical distribution of $\hat{t}_j^{p=1}, \ldots \hat{t}_j^{p=1000}$ to calculate the p-value associated with $\hat{t}_j$ for each j.

[16] Y_{0iv} will be the baseline value of the second-stage outcome for all first stage models, as the baseline value of the first stage outcomes are zero for all households.

[17] A separate paper will compare the treatment effects of psychological interventions that are delivered before, and ones that are delivered after, cash transfers. See https://www.socialscienceregistry.org/docs/analysisplan/807/document for details.

We will implement these additional hypothesis tests only for the five primary outcomes indices, and only when testing whether the coefficients $(\beta_C, \beta_P, \beta_{CP})$ are individually equal to zero. So we will not account for multiple testing when implementing the bootstrap and permutation tests. We will calculate the indices within each iteration, which will account for the covariance between the components of the index.

9 Subgroup 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 baseline value of the outcome in each model 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 respondent's education level is above the sample median.
4. An indicator variable equal to one if respondent's household income is above the sample median.
5. An indicator variable equal to one if respondent's asset index (defined in section 4) is above the sample median.
6. An indicator variable equal to one if respondent is unmarried or widowed.
7. An indicator variable equal to one if the age of the child about whom the education questions are asked is above the sample median (only for variables ed_1, ed_2, and ed_3).

For non-binary outcomes, we will also estimate quantile treatment effects using an estimation method adapted from Firpo (2007). We use the model:

$$Q\left(\tilde{Y}_{iv}\right) = \beta_0^Q + \text{Cash}_v \cdot \beta_C^Q + \text{Psych}_v \cdot \beta_P^Q + \text{Cash}_v \cdot \text{Psych}_v \cdot \beta_{CP}^Q + \epsilon_{iv}, \tag{3}$$

where $Q\left(\tilde{Y}_{iv}\right)$ is the Q^{th} quantile of the distribution of $Y_{iv} \cdot \omega\left(G_v; Y_{0iv}, T_{iv}, \alpha_v\right)$, G is a vector of treatment assignment indicators, and

$$\omega\left(g; Y_{0iv}, T_{ivP}, T_{ivC}, \alpha_v\right) = 1/Pr\left(G_v = g | Y_{0iv}, T_{iv}, \alpha_v\right) \tag{4}$$

is a weight function. We will estimate the weights by regressing a vector of treatment assignments on $(Y_{0iv}, T_{iv}, T_{iv}^2, \alpha_v)$ using a multinomial logit, constructing the predicted probabilities, winsorizing the top and bottom percentiles, and normalizing the weights to sum to one within each treatment group. We will construct the covariance matrix for $\left(\beta_0^Q, \beta_C^Q, \beta_P^Q, \beta_{CP}^Q\right)$ at each Q by bootstrapping the entire estimation process (logit regression, prediction, winsorizing, normalizing, quantile regression).

We will also estimate the model

$$Y_{iv} = \text{Psych}_v \cdot \tilde{\beta}_P + Y_{0iv} \cdot \gamma_0 + T_{iv} \cdot \tau_1 + T_{iv}^2 \cdot \tau_2 + \alpha_v + \epsilon_{iv}, \tag{5}$$

using only households who were among the first batch of Lump Sum A transfer recipients in their village. This omits households that were not assigned to receive cash transfers, households that consistently refused cash transfers, and households that initially refused transfers but changed their mind when the psychological intervention was already underway. $\tilde{\beta}_P$ 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.

10 Analysis of Group Composition Effects

We will also test if the randomly assigned composition of intervention groups in Siaya affects outcomes. As treatment is assigned at the village level, there is no variation in assigned treatment within intervention groups. So this analysis examines only effects of peer group composition, not spillover effects of treatment status.

We will estimate the standard linear-in-means peer effects model

$$Y_{igsv} = \alpha_{sv} + \overline{Y}_{-igvs} \cdot \beta + \epsilon_{igsv} \tag{6}$$

where $\overline{Y}_{-igvs}$ is the mean value of Y for individuals in group g excluding individual i and α_{sv} is a village-session fixed effect. We will estimate this model for each outcome Y listed in section 3. We will also estimate

$$Y_{igsv} = \alpha_{sv} + \overline{Y}_{-igvs}^{linked} \cdot \gamma + \overline{Y}_{-igvs}^{unlinked} \cdot \delta + \epsilon_{igsv} \tag{7}$$

where $\overline{Y}^{linked}_{-igvs}$ is the mean value of Y for individuals in group g excluding individual i and individuals i did not know before the intervention and $\overline{Y}^{unlinked}_{-igvs}$ is the mean value of Y for individuals in group g excluding individual i and individuals i did know before the intervention. We ask all members of a group about their relationship with other group members before starting the intervention.

We will report standard errors clustered at the group level. We will also report p-values for the null hypotheses of:

1. no peer effects in equation (6)
2. no linked-peer effects in equation (7)
3. no unlinked-peer effects in equation (7)
4. equal linked-peer or unlinked-peer effects in equation (7)
5. no linked-peer or unlinked-peer effects in equation (7)

obtained by permuting the group assignments within each village-session.

References

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, 1481–1495.

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

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*, 1948550613476307.

Dweck, C. S. (2012): "Mindsets and human nature: Promoting change in the Middle East, the schoolyard, the racial divide, and willpower," *American Psychologist*, 67, 614–622.

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

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

Kling, J., J. Liebman, and L. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75, 83–119.

Lamping, D. L., S. Schroter, P. Marquis, A. Marrel, I. Duprat-Lomon, and P.-P. Sagnier (2002): "The community-acquired pneumonia symptom questionnaire: a new, patient-based outcome measure to evaluate symptoms in patients with community-acquired pneumonia," *CHEST Journal*, 122, 920–929.

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

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, 255–264.

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

Schwarzer, R. and M. Jerusalem (1995): *Generalized Self-Efficacy scale*, Windsor, UK: NFER-NELSON.

Shao, J. and R. Sitter (1996): "Bootstrap for Imputed Survey Data," *Journal of the American Statistical Association*, 91, 1278–1288.

Appendix

A Group size rules

# people	group 1	group 2	rescheduled
1			1
2	2		
3	3		
4	4		
5	3	2	
6	4	2	
7	4	3	
8	4	4	
9	4	4	1
10	4	4	2
11	4	4	3
12	4	4	4
13	4	4	5
14	4	4	6
15	4	4	7
16	4	4	8

Table 1: Two groups (morning session)

# people	group 1	group 2	rescheduled
1			1
2	2		
3	3		
4	4		
5	3	2	
6	4	2	
7	4	3	
8	4	4	
9	5	4	
10	6	4	
11	6	5	
12	6	6	
13	4	4	5
14	4	4	6
15	4	4	7
16	4	4	8

Table 2: Two groups (evening session)

# people	group 1	group 2
1	1	
2	2	
3	3	
4	4	
5	3	2
6	4	2
7	4	3
8	4	4
9	5	4
10	6	4
11	6	5
12	6	6
13	7	6
14	7	7
15	8	7
16	8	8

Table 3: Two groups without rescheduling

# people	group 1	group 2	group 3	rescheduling
1				1
2	2			
3	3			
4	4			
5	3	2		
6	4	2		
7	4	3		
8	4	4		
9	3	3	3	
10	4	3	3	
11	4	4	3	
12	4	4	4	
13	4	4	4	1
14	4	4	4	2
15	4	4	4	3
16	4	4	4	4
17	4	4	4	5
18	4	4	4	6
19	4	4	4	7
20	4	4	4	8

Table 4: Three groups (morning session)

# people	group 1	group 2	group 3	rescheduling
1				1
2	2			
3	3			
4	4			
5	3	2		
6	4	2		
7	4	3		
8	4	4		
9	3	3	3	
10	4	3	3	
11	4	4	3	
12	4	4	4	
13	5	4	4	
14	6	4	4	
15	6	5	4	
16	6	6	4	
17	4	4	4	5
18	4	4	4	6
19	4	4	4	7
20	4	4	4	8

Table 5: Three groups (evening session)

# people	group 1	group 2	group 3
1	1		
2	2		
3	3		
4	4		
5	3	2	
6	4	2	
7	4	3	
8	4	4	
9	3	3	3
10	4	3	3
11	4	4	3
12	4	4	4
13	6	4	3
14	6	4	4
15	6	5	4
16	6	5	5
17	6	6	5
18	6	6	6
19	7	6	6
20	8	6	6

Table 6: Three groups without rescheduling