

Welfare Effects of Unconditional Cash Transfers: Pre-Analysis Plan*

Johannes Haushofer † Jeremy Shapiro ‡

March 10, 2016

Abstract

This document describes the analysis plan for a randomized controlled trial (RCT) evaluating the impacts of unconditional cash transfers. We outline analysis planned for a second endline survey, building on our previous analysis plan (dated June 27, 2013) and analysis relating to the initial endline survey. We study the effect of unconditional cash transfers distributed in Kenya by the NGO GiveDirectly between June 2011 and January 2013. GiveDirectly distributed cash transfers to 503 randomly selected rural households in Western Kenya. Transfers were sent to recipients' mobile phones using the M-Pesa technology. The RCT included three sub-treatment arms: First, the transfers were randomly chosen to be sent to either the primary female or the primary male member of the household. Second, the transfers were randomly assigned to be sent as either a large lump-sum payment, or a series of nine monthly installments of the same total amount. Third, the magnitude of the total transfer to each treatment household was randomly chosen to be either USD 404 PPP or USD 1,525 PPP. An initial endline was conducted in 2013, on average nine months after the beginning of transfers. That data was analyzed in an earlier study on the short run impacts of cash transfers. This document outlines the outcome variables and econometric methods we will use to assess the longer run effect of cash transfers, approximately three years after beginning of transfers.

*We thank Jim Reisinger, Catherine Thomas, Faizan Diwan, Channing Jang, Chris Roth, Alexis Grigorieff, and James Vancel for excellent research assistance, the team of GiveDirectly for collaboration, and Petra Persson for designing the intrahousehold bargaining and domestic violence module. This research was supported by NIH Grant R01AG039297 to Johannes Haushofer and by a grant from an anonymous donor to Jeremy Shapiro and Johannes Haushofer. The data reported in this paper will be combined with those from a further long-term follow-up on GiveDirectly unconditional cash transfers in 2017/2018, in joint work with Ted Miguel, Paul Niehaus, and Michael Walker.

†Princeton University Haushofer@princeton.edu

‡Busara Center for Behavioral Economics jeremy.shapiro@busaracenter.org

JEL Codes: C93, D13, I15, I25, O12

Keywords: unconditional cash transfers, randomized controlled trial, impact evaluation.

1 Introduction

Cash transfers are a simple intervention to improve the lives of the poor, and have attracted attention from governments, donors and researchers . Often these transfers have been in-kind (Brune et al. 2015; Bandiera et al. 2013; Banerjee et al. 2015), or conditional (Maluccio, Murphy, and Regalia 2010; Maluccio and Flores 2005; Barham, Macours, and Maluccio 2013; Maluccio 2010; Baird, McIntosh, and Özler 2011; Baird et al. 2012; Baird, De Hoop, and Özler 2013; Filmer and Schady 2009; Ferreira, Filmer, and Schady 2009; Skoufias, Unar, and Cossio 2013; Akresh, de Walque, and Kazianga 2013; Macours, Schady, and Vakis 2012; Attanasio and Mesnard 2006; Sadoulet, Janvry, and Davis 2001; Gertler, Martinez, and Rubio-Codina 2012). Recently, unconditional cash transfers have gained widespread attention (Aguëro, Carter, and Woolard 2010; Aker et al. 2016; Aker 2015; Amarante et al. 2010; Ardington, Case, and Hosegood 2009; Baird, McIntosh, and Özler 2011; Baird et al. 2012; Baird, De Hoop, and Özler 2013; Baird et al. 2016; Blattman, Fiala, and Martinez 2014; Blattman, Jamison, and Sheridan 2015; Cunha 2014; Currie and Gahvari 2008; De Mel, McKenzie, and Woodruff 2008; Duflo 2003; Edmonds and Schady 2012; Edmonds 2006; Fafchamps and Quinn 2015; Fafchamps et al. 2011; Fernald and Hidrobo 2011; Mckenzie 2015; Oosterbeek, Ponce, and Schady 2008; Paxson and Schady 2010; Posel, Fairburn, and Lund 2006; Schady and Rosero 2008; Schady and Araujo 2008; Team 2012a; Team 2012b).

Our initial study was designed to contribute to the evidence on the impact of unconditional cash transfers on economic and psychological outcomes. We collaborated with the NGO GiveDirectly to study the impact of cash transfers in Rarieda, Western Kenya. In late 2013, we conducted an endline survey – approximately nine months after the beginning of the transfers to study households. The analysis of that data is documented in a summary policy brief, and a full treatment is given in a paper currently under review. One limitation of the initial analysis was the relatively short time horizon between the end of the transfers and outcome data collection. We therefore conducted a second endline survey in 2015. In analyzing the second endline data, we aim to answer the following questions:

1. What are the long-run impacts of broadly targeted unconditional cash transfers on economic outcomes, psychological well-being, IPV, and on non-recipients?
2. How do long-run impacts vary based on design parameters (transfer frequency, recipient gender, and transfer magnitude)?
3. What are potential mechanisms generating the reduction in IPV in the treatment and spillover sample (documented in the initial analysis)?

Below we outline our approach to answering these questions. This document was written before the commencement of analysis of the data. Generally we follow the approach laid out in our previous analysis plan (initial and amended with supplementary analysis) with some modifications based on updated hypotheses, survey instruments, and changes to the sampling design.

2 Intervention

A goal of this study was to assess the relative impacts of three design features of unconditional cash transfers on economic and other outcomes, to assess whether households effectively pool income, whether they are credit- and savings-constrained, and how the magnitude of transfers affects outcomes. To address these questions, we randomized the gender of the transfer recipient, the temporal structure of the transfers (monthly vs. lump-sum transfers), and the magnitude of the transfer. The treatment arms were structured as follows:

1. **Transfers to the woman vs. the man in the household.** Among households with both a primary female and a primary male member, we stratified on recipient gender and randomly assigned the woman or the man to be the transfer recipient with equal probability. 110 households had a single household head and were not considered in the randomization of recipient gender.
2. **Lump-sum transfers vs. monthly installments.** Across all treatment households, we randomly assigned the transfer to be delivered either as a lump-sum amount or as a series of nine monthly installments. Specifically, 258 of the 503 treatment households were assigned to the monthly condition, and 245 to the lump-sum condition. In the analysis we only consider the 173 monthly recipient and 193 lump-sum recipient households that did not receive large transfers, because large transfers were not unambiguously monthly or lump-sum (see below). The total amount of each type of transfer was KES 25,200 (USD 404 PPP). In the lump-sum condition, this amount includes an initial transfer of KES 1,200 (USD 19 PPP) to incentivize *M-Pesa* registration, followed by a lump-sum payment of KES 24,000 (USD 384 PPP). In the monthly condition, the total amount consists of a sequence of nine monthly transfers of KES 2,800 (USD 45 PPP) each. The timing of transfers was structured as follows: In the monthly condition, recipients received the first transfer of KES 2,800 on the first of the month following *M-Pesa* registration, and the remaining eight transfers of KES 2,800 on the first of the eight following months. In the lump-sum condition, recipients received the initial transfer of KES 1,200 on the first of the month following *M-Pesa*

registration, and the lump-sum transfer of KES 24,000 on the first of a month that was chosen randomly among the nine months following the time at which they were enrolled in the *GD* program.

3. **Large vs. small transfers.** Finally, a third pair of treatment arms was created to study the relative impact of large compared to small transfers. To this end, 137 households in the treatment group were randomly chosen and informed in January 2012 that they would receive an additional transfer of KES 70,000 (USD 1,121 PPP), paid in seven monthly installments of KES 10,000 (USD 160 PPP) each, beginning in February 2012. Thus, the transfers previously assigned to these households, whether monthly or lump-sum, were augmented by KES 10,000 from February 2012 to August 2012,¹ and therefore the total transfer amount received by these households was KES 95,200 (USD 1,525 PPP, USD 1,000 nominal). The remaining 366 treatment households constitute the “small” transfer group, and received transfers totaling KES 25,200 (USD 404 PPP, USD 300 nominal) per household.

These three treatment arms were fully cross-randomized, except that, as noted above, the “large” transfers were made to existing recipients of KES 25,200 transfers in the form of a KES 70,000 top-up that was delivered as a stream of payments after respondents had already been told that they would receive KES 25,200 transfers.

3 Evaluation Design

3.1 Sampling and Identification strategy

This study is a two-level cluster-randomized controlled trial. The selection and surveying of recipient households proceeded as follows:

1. *GD* first identified Rarieda, Kenya, as a study district, based on data from the national census. The research team then identified the 120 villages with the highest proportion of thatched roofs within Rarieda. Sixty villages were randomly chosen to be treatment villages (first stage of randomization). Villages had an average of 100 households. An average of 19 percent of households per village were surveyed, and an average of 9 percent received transfers. The transfers sent to villages amounted to an average of 10 percent of aggregate baseline village wealth.

¹Note that for the households originally assigned to the “lump-sum” condition, this new transfer schedule implied that these households could no longer be unambiguously considered to be lump-sum households; we therefore restrict the comparison of lump-sum to monthly households to those households which received small transfers, as described above.

2. The research team then identified all eligible households within treatment villages through a census administered with the assistance of the village elder. Census exercises were conducted in March–November 2011 in treatment villages, and in April–June 2012 in control villages. The census was conducted in the same fashion in treatment and control villages. A household was considered eligible if it had a thatched roof. The purpose of the census and baseline was described to village elders and respondents as providing information to researchers about living conditions in the area; no mention was made of *GD* or transfers.
3. Following the census, all eligible households completed the baseline survey. Thus, baseline surveys were administered between April and November 2011. The order of census and surveys was randomized at the village level (after the first four villages, which were chosen for proximity to the field office). No transfers or transfer announcements were made before or during census or baseline in each village. The surveys were described to respondents in the same fashion as the census, that is, without reference to *GD* or transfers.
4. *GD* then repeated the census to confirm that all households deemed eligible by the research team were in fact eligible. The final eligible sample was the overlap between the households that completed baseline and *GD*'s census exercise. We excluded 89 households who completed baseline but were not identified as eligible in the *GD* census. After baseline, the research team randomly chose half of the eligible households to be transfer recipients (second stage of randomization). This process resulted in 503 treatment households and 505 control households in treatment villages at baseline. We refer to the control households in treatment villages as “spillover” households.
5. Within a few weeks after all households in a village had completed baseline and the *GD* census, recipient households were visited by a representative of *GD*, who announced the transfer, including amount and timing (although note that large transfers were announced later as a top-up to existing small transfers). We have no data on how transfers were perceived by the households; anecdotally, because *GD* worked with village elders, had objectively verifiable targeting criteria, and was otherwise highly transparent, we have reason to believe that recipients had accurate beliefs about the nature of the transfers as fully unconditional and one-time. Control households were not visited, but those who asked were told that they had not won the lottery for transfers. The control group did not receive SIM cards and were not asked to register for *M-Pesa*; thus, our treatment effects reflect the joint impact of cash transfers and incentives to register for *M-Pesa* (Jack and Suri 2014).

6. The transfer schedule commenced on the first day of the month following the initial visit. For monthly transfers, the first installment was transferred on that day, and continued for eight months thereafter; for lump-sum transfers, a month was randomly chosen among the nine months following the date of the initial visit. Each transfer was announced with a text message; recipients who did not own cell phones could rely on the transfer schedule given to them by *GD* to know when they would receive transfers, or insert the SIM card into any mobile handset periodically to check for incoming transfers. To facilitate transfer delivery, *GD* offered to sell cell phones to recipient households which did not own one (by reducing the future transfer by the cost of the phone).
7. An endline survey was administered by the research team between August and December 2012. The order in which villages were surveyed followed the same order as the baseline. In a small number of households, the endline survey was administered before the final transfer was received. These households are nevertheless included in the analysis to be conservative (intent-to-treat). Control villages were surveyed only at endline; in these villages, we sampled 432 households from among eligible households. We refer to these households “pure control” households. The census exercise to select these households was identical to that in treatment villages, except that no *GD* census was administered. Because these pure control households were selected into the sample just before the endline, the thatched-roof criterion was applied to them about one year later than to households in treatment villages. This fact potentially introduces bias into the comparison of households in treatment and control villages; we describe below how it was dealt with in the second endline.
8. We administered a second endline between February and September of 2015. In the second endline survey we made two changes to the sampling design. First, we surveyed an additional 349 households in pure control villages. The comparison of these households to previously surveyed households in pure control villages identifies demand effects. The analysis to identify these demand effects will be covered in a separate pre-analysis plan. Second, we surveyed 71 additional households in the pure control villages to correct for the fact that the eligibility criteria (living in a thatched roof house) was applied one year later in pure control villages than in treatment villages (this is discussed below and in the analysis of the first endline).

This study will primarily focus on the results of this second endline.

3.2 Power calculation

Using data from the initial endline survey, we determined that for our core specification of within-village treatment effects, we can detect the following impacts with 80% power: a 15% change in asset holdings, a 10% change in non-durable expenditure, a 0.17-0.20 SD change in standardized indices, and a 34% change in business and agricultural revenue.

3.3 Risk and treatment of attrition

In treatment villages, attrition between baseline and the first endline was 6.3% among treatment households and 7.1% among control households. Between the baseline and follow-up, attrition was 9.1% for treatment households and 9.3% for control households.

In control villages, attrition between the first endline and follow-up was 13%. We only report attrition between the first endline and follow-up, as there was no baseline survey. We discuss our approach for dealing with attrition econometrically below.

4 Econometric specifications

4.1 Baseline Balance

In previous analysis we tested for baseline differences between treatment and control groups using the following specification:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhiB} \quad (1)$$

Here, y_{vhiB} is the outcome of interest for household h in village v , measured at baseline, of individual i (subscript i is included for outcomes measured at the level of the individual respondent, and omitted for outcomes measured at the household level). To maximize power, and because pure control villages were not surveyed at baseline, we focus on the within-village treatment effect and therefore restrict the sample to treatment and control households in treatment villages.² Village-level fixed effects are captured by α_v . T_{vh} is a treatment indicator that takes value 1 for treatment households, and 0 otherwise. ε_{vhiB} is an idiosyncratic error term. The omitted category is control households in treatment villages; thus, β_1 identifies the difference in baseline outcomes between treated households and control households in treatment villages. Standard errors are clustered at the level of the unit of randomization,

²A further reason to focus on the within-village treatment effect is that the within-village treatment and control groups participated in the same number of surveys, minimizing concerns about survey effects (Zwane et al. 2011; Baird and Özler 2012).

that is, the household. For individual-level outcomes, we use inverse probability weights to weight households and villages equally. In addition to this standard inference, we compute FWER-corrected p -values across the set of index variables. Finally, we estimate the system of equations jointly using seemingly unrelated regression (SUR), which allows us to perform Wald tests of joint significance of the treatment coefficient.

We also analyzed the difference in baseline outcomes between treatment arms, restricting the sample to treatment villages. To calculate differences between treatment households in which transfers were made to the female vs. the male in the household, we used the following specification:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^F + \beta_2 T_{vh}^W + \beta_3 S_{vh} + \varepsilon_{vhiB} \quad (2)$$

Here, the variables T_{vh}^x are indicator functions that specify whether the transfer recipient is female (T_{vh}^F) or that the gender of the recipient could not be randomized because the household had only one head (most commonly in the case of widows/widowers) (T_{vh}^W). S_{vh} is an indicator variable for the spillover group. The omitted category is two-headed households in which the primary male received a transfer. β_1 is the difference in baseline outcomes between female and male recipient households.

To assess baseline differences between monthly vs. lump-sum transfers, we analyzed the following:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{MTH} \times T_{vh}^S + \beta_2 T_{vh}^L + \beta_3 S_{vh} + \varepsilon_{vhiB} \quad (3)$$

Here, T_{vh}^{MTH} is an indicator variable for having been assigned to monthly transfers, and T_{vh}^S and T_{vh}^L for being assigned to the small and large transfer conditions, respectively. Note that households assigned to the large transfer condition cannot unambiguously be considered monthly or lump-sum, and therefore this regression compares households which did not receive large transfers. The omitted category is thus households that received a (small) lump-sum transfer. β_1 is the difference in baseline outcomes between monthly and lump-sum recipient households.

Finally, to assess baseline differences between households receiving large compared to small transfers, we used the following specification:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^L + \beta_2 S_{vh} + \varepsilon_{vhiB} \quad (4)$$

Here, T_{vh}^L is an indicator variable for having been assigned to receiving large transfers. Thus, β_1 is the difference in baseline outcome measures between households receiving large transfers and households receiving small transfers.

4.2 Reduced form specifications

Our basic treatment effects specification to capture the impact of cash transfers is

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (5)$$

Here, y_{vhiE} is the outcome of interest for household h in village v , measured at endline, of individual i (subscript i is included for outcomes measured at the level of the individual respondent, and omitted for outcomes measured at the household level). As for the analysis of baseline balance, we again restrict the sample to treatment and control households in treatment villages; we discuss bounds for the spillover effect in Section 4.3. Following McKenzie (2012), we condition on the baseline level of the outcome variable when available, y_{vhiB} , to improve statistical power. To include observations where the baseline outcome is missing, we code missing values as 0 and include a dummy indicator that the variable is missing (M_{vhiB}). All other features are as in Equation 1.

To distinguish between the effects of different treatment arms, we use the analogous versions of Equations 2, 3, and 4. First, the effect of making the transfer to the female vs. the male in the household is captured by the following model, restricting the sample to treatment villages and denoting spillover households with S_{vh} :

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^F + \beta_2 T_{vh}^W + \beta_3 S_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (6)$$

The specification to assess the relative effect of monthly vs. lump-sum transfers is as follows:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{MTH} \times T_{vh}^S + \beta_2 T_{vh}^L + \beta_3 S_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (7)$$

Finally, the specification to assess the effect of receiving large compared to small transfers is

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^L + \beta_2 S_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (8)$$

All restrictions and other features are as described in Section 4.1.

4.3 Estimating spillover effects

For the within-village specifications to provide an unbiased estimate of the treatment effect, within-village spillovers of treatment on non-recipient households must be small. This includes both spillover effects that operate through economic channels, and those that have psychological roots, such as John Henry effects. To address this question, we estimate the

magnitude of these within-village spillovers by comparing spillover to pure control households:

$$y_{vhiE} = \beta_0 + \beta_1 S_{vh} + \varepsilon_{vhiE} \quad (9)$$

Here, the sample includes only non-treatment households (in treatment and control villages). Thus, β_1 identifies within-village spillover effects by comparing control households in treatment villages to control households in pure control villages. The error term is clustered at the village level, reflecting the dual-level randomization at the village and within-village (household) levels (Cameron, Gelbach, and Miller 2011; Pepper 2002). Note that the inclusion of baseline covariates is not feasible here because no baseline data exist for the pure control group. Similarly, village-level fixed effects are not feasible because they would be collinear with S_{vh} .

A potential weakness in the spillover analysis is that the thatched-roof selection criterion for participation in the study was applied to households in control villages one year after it was applied to households in treatment villages. As a result, there is endogenous selection into the pure control condition, as some proportion of households in pure control villages are likely to have upgraded to a metal roof over this time period. These households are excluded from endline in the pure control villages, potentially introducing bias into the spillover analysis.

To deal with this potential bias, as part of the follow-up survey, we visited households in pure control villages that purchased metal roofs between the dates of the baseline and first endline surveys. Since these are the households that were excluded due to the late application of the thatched-roof selection criterion, including them allows us to calculate unbiased spillover effect estimates as of the follow-up. Our approach was as follows:

1. We first assessed the reliability with which individuals could recall the date when they upgraded their roof. To do this, we asked households who upgraded in treatment villages when they upgraded their roof. Since we have objective data indicating whether they upgraded between baseline and endline, we can assess the proportion who accurately place the date in or out of that window. We surveyed 108 households we know upgraded to a metal roof between baseline and endline surveys (from our objective data), of those: 78 respondents (72.2%) reported upgrading within the baseline and endline 1 window, 17 respondents (15.7%) reported upgrading outside the baseline and endline 1 window, 13 respondent (12.0%) could not recall at all.
2. Having established reasonable reliability of the date of recall, we returned to all households with iron roofs in pure control villages to inquire when they upgraded their roof.

If they informed us that they upgraded at a date between baseline and endline 1, we classify them as eligible to be surveyed as part of the pure control group (though they were excluded in endline 1). We refer to these households as “newly eligible pure control” (NEPC), in contrast to originally eligible pure control households (OEPC). We determined there were 170 NEPC households during this exercise.

3. We then used the same algorithm originally used to select pure control households to calculate the probability that each of these households would have been included in the study had they been identified as eligible at the time. The original sampling method required us to select 8 households from the pool of eligible households in each village (those with thatched roofs). When there were 8 or fewer eligible households in a given village, we selected all households. When there more than 8 eligible households, we selected 8 with equal probability for each. We were thus able to calculate the exact probability that a given household would be selected in each village. In villages with 8 or fewer eligible households, the probability of selection was 1. In villages with more than 8 eligible households, the probability was 8 divided by the total number of eligible households.
4. Based on these probabilities, we then selected a subsample of 71 NEPC households to survey at the follow-up that we include in this analysis. We will include these additional 71 households in all analyses involving pure control households. As a robustness check, we will also report results using the methods employed in our analysis of the endline 1 data.

4.4 Across-village comparison

In the presence of spillover effects, our treatment effect estimates in Section 4.2 are biased. We therefore also estimate the treatment effect of cash transfers by comparing treatment households to pure control households. In this analysis, we omit spillover households. The specification is

$$y_{vhiE} = \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhiE} \quad (10)$$

where variables are as defined above. We omit village-level fixed effects, as they would be collinear with treatment status. Additionally, since pure control households were not surveyed at baseline, we do not include baseline values of outcome variables on the right-hand side. Since this analysis leverages village-level randomization, we cluster standard errors at the village level. Thus β_1 identifies the treatment effect relative to control households in

control villages. Any difference between the treatment effect measured in this specification and the within-village specification will be due to spillover effects, an issue examined in more depth using the analysis outlined in Section 4.3.

We also analyze the various treatment arms using across-village comparisons using analogous versions of Equations 6, 7, and 8. For across-village treatment effect for households in which the primary male and primary female received the transfer, we now include pure control households in the analysis and estimate:

$$y_{vhiE} = \beta_0 + \beta_1 T_{vh}^F + \beta_2 T_{vh}^M + \beta_3 T_{vh}^W + \beta_4 S_{vh} + \beta_5 PC_{vh}^{\text{SINGLE}} + \varepsilon_{vhiE} \quad (11)$$

Here, PC_{vh}^{SINGLE} is an indicator for pure control households with a single head. Thus, the omitted category is cohabiting pure control households. β_1 identifies the treatment effect when the primary female in the household receives the transfer. β_2 identifies the treatment effect when the primary male in the household receives the transfer. Again we omit baseline outcomes, as they were not measured for households in the pure control group, and cluster standard errors at the village level.

For across-village treatment effect for monthly and lump-sum transfers, we now include pure control households and estimate:

$$y_{vhiE} = \beta_0 + \beta_1 T_{vh}^{\text{MTH}} \times T_{vh}^S + \beta_2 T_{vh}^{\text{LS}} \times T_{vh}^S + \beta_3 T_{vh}^L + \beta_4 S_{vh} + \varepsilon_{vhiE} \quad (12)$$

Thus, the omitted category is pure control households. β_1 identifies the effect of a monthly transfer using an across-village comparison. β_2 identifies the effect of a lump-sum transfer using an across-village comparison. Again we omit baseline outcomes, as they were not measured for households in the pure control group, and cluster standard errors at the village level.

For across-village treatment effect for large and small transfers, we include pure control households and estimate:

$$y_{vhiE} = \beta_0 + \beta_1 T_{vh}^L + \beta_2 T_{vh}^S + \beta_3 S_{vh} + \varepsilon_{vhiE} \quad (13)$$

Thus, the omitted category is pure control households. β_1 identifies the effect of a large transfer using an across-village comparison. β_2 identifies the effect of a small transfer using an across-village comparison. Again we omit baseline outcomes and cluster standard errors at the village level.

4.5 Comparison of short and long-run impacts

Since we have outcome data measured in the short run (~9 months after the beginning of the transfers) and in the long-run (~3 years after the beginning of transfers), we will test equality between short and long-run effects. Specifically, we will combine endline one and endline two data in long form and estimate:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh} \times E_1 + \beta_2 T_{vh} \times E_2 + \beta_4 E_1 + \beta_5 E_2 + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (14)$$

Here, y_{vhiE} is the outcome of interest for household h in village v , measured at endline one or two, of individual i (subscript i is included for outcomes measured at the level of the individual respondent, and omitted for outcomes measured at the household level). The sample is restricted to treatment villages. T_{vh} is a treatment indicator that takes value 1 for households which received a cash transfer (“treatment households”) and 0 otherwise. E_1 and E_2 take value one if the data in the row is from endline one or two, respectively. A test of $\beta_1 = \beta_2$ allows us to establish whether impacts vary over time. Analogues of this equation will be run for the across-village estimates, treatment arms and spillover effects, interacting the relevant treatment variables with E_1 and E_2 .

4.6 Additional analyses

We are primarily interested in the main effect of cash transfers, spillover effects, and variation in effects based on program design (treatment arm) – the analyses described above. In addition, we will also explore heterogeneous effects, and we outline our approach to doing so below.

4.6.1 Comparison of recipient vs. spouse within households

In most analyses, we consider the individuals within a household to be “treated” if that household received a transfer. However, there may be differences in effect when an individual receives the transfer himself or herself and when an individual’s spouse receives the transfer. We will not be able to detect this difference for outcomes measured at the household level, but for outcomes measured at the individual level, the relevant specification takes the following form, restricting to treatment villages:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\text{SELF}} + \beta_2 T_{vh}^{\text{SINGLE}} + \beta_3 S_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (15)$$

In this specification, T_{vh}^{SELF} is an indicator variable which takes value 1 for respondents who live with a spouse and received the transfer themselves, and 0 otherwise. T_{vh}^{SINGLE} is an indicator for an individual who received a transfer and does not live with a spouse. The omitted category is T_{vh}^{SPOUSE} , for individuals in treatment households whose spouse received the transfer. Thus, in this model, β_1 identifies the difference in treatment effect for cohabiting individuals when the individual receives the transfer and when their spouse receives the transfer.

For the across-village treatment effect for households depending on whether the respondent or his / her spouse received the transfer, we include pure control households and estimate:

$$y_{vhiE} = \beta_0 + \beta_1 T_{vh}^{SELF} + \beta_2 T_{vh}^{SPOUSE} + \beta_3 T_{vh}^{SINGLE} + \beta_4 S_{vh}^{COUPLE} + \beta_5 S_{vh}^{SINGLE} + \beta_6 P_{vh}^{SINGLE} + \varepsilon_{vhiE} \quad (16)$$

Here, P_{vh}^{SINGLE} is an indicator variable taking the value of 1 for single-headed pure control households. Thus, the omitted category is P_{vh}^{COUPLE} , cohabiting pure control households. β_1 is the across-village effect when the respondent receives the transfer. β_2 is the across-village effect when the respondent's spouse receives the transfer. Standard errors are clustered at the village level.

4.6.2 Heterogeneous treatment effects

We will test whether the impact of the cash transfers varies with pre-determined household and individual characteristics, measured at baseline and denoted by X , which may be either household- or individual-level characteristics. This allows us to understand which sub-groups contribute disproportionately to overall treatment effects. Note that the pure control group was not included in the baseline; we will therefore estimate, for household-level outcomes:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{hv} + \beta_2 X_{vhi} + \beta_3 T_{vh} \times X_{vhi} + \delta y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE} \quad (17)$$

As before, standard errors in this type of specification will be clustered at the household level, and the sample is restricted to treatment villages.

We will also estimate analogues of this specification which incorporate the main effects and all interaction terms of additional treatment arms (gender or recipient, frequency of transfer, and magnitude of the transfer).

These analyses will take the form following forms:

$$\begin{aligned}
y_{vhiE} = & \alpha_v + \beta_0 + \beta_1 T_{vh}^F + \beta_2 T_{vh}^M + \beta_3 T_{vh}^W + \beta_4 T_{vh}^F \times X_{vhi} + \beta_5 T_{vh}^M \times X_{vhi} \\
& + \beta_6 T_{vh}^W \times X_{vhi} + \beta_8 X_{vhi} + \beta_9 S_{vh}^{\text{SINGLE}} + \delta y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE}
\end{aligned} \tag{18}$$

$$\begin{aligned}
y_{vhiE} = & \alpha_v + \beta_0 + \beta_1 T_{vh}^{\text{MTH}} \times T_{vh}^S + \beta_2 T_{vh}^{\text{LS}} \times T_{vh}^S + \beta_3 T_{vh}^{\text{MTH}} \times T_{vh}^S \times X_{vhi} \\
& + \beta_4 T_{vh}^{\text{LS}} \times T_{vh}^S \times X_{vhi} + \beta_5 X_{vhi} + \beta_6 T_{vh}^L + \delta y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE}
\end{aligned} \tag{19}$$

$$\begin{aligned}
y_{vhiE} = & \alpha_v + \beta_0 + \beta_1 T_{vh}^L + \beta_2 T_{vh}^S + \beta_3 T_{vh}^L \times X_{vhi} + \beta_4 T_{vh}^S \times X_{vhi} \\
& + \beta_5 X_{vhi} + \delta y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhiE}
\end{aligned} \tag{20}$$

Again all restrictions and other features are as described in Section 4.1. Individual outcomes may be tested for heterogeneous impacts both in terms of variables which vary at the household level, X_{vh} , as well as those that vary at the individual level, X_{vhi} . In cases where the outcome variable is at the household level and heterogeneity is at the individual level, the value of the individual-level variable for the transfer recipient (rather than their spouse) will be used as the source of heterogeneity.

Dimensions of heterogeneous effects

Heterogeneous effects will be considered along the following dimensions:

1. Age of primary household members
2. Baseline consumption, food security, asset levels, and land holdings (index)
3. Psychological wellbeing (index)
4. Center for Epidemiological Studies - Depression scale (3 level k-means cluster) - in earlier work, we defined 3 “mindset” clusters based on this scale. We will explore heterogeneous effects across these groups.

4.6.3 Other extensions

To further our understanding of the long-term effects of cash transfers, we anticipate combining this data with data collected in an ongoing study by Haushofer, Miguel, Niehaus, and Walker. Analysis of combined data will be covered in a separate pre-analysis plan.

4.7 IPV Analysis

We plan to test a number of exploratory hypotheses related to IPV reduction, and also to analyze extensive qualitative data collected alongside the second endline survey.

4.7.1 IPV hypotheses

We formulated a number of hypotheses related to IPV in our qualitative work, we will test these hypotheses using the specifications described to evaluate heterogeneous effects.

1. Index of intrahousehold bargaining and domestic violence variables (described below) – we predict those with initially higher levels of IPV will exhibit the greatest reductions
2. Duration of cohabitation with marital partner – we predict that more recently married and not married but cohabiting couples will report more IPV at baseline and show larger treatment effects on IPV compared to other couples
3. Stressful life events, including pregnancy – we predict that households experiencing an above median number of stressful life events, such as pregnancy, in the past year will report more IPV and attenuated treatment effects on IPV compared to others those experiencing a below median number of stressful life events
4. Education levels – we predict less IPV in a relationship and attenuated treatment effects in households in which the female and/or the male respondent have at least secondary education and the spouses have small to no disparity in their education levels
5. Self-Esteem (Rosenberg) – we predict less IPV in a relationship and attenuated treatment effects in households in which the female and/or the male respondent exhibit high self-esteem at baseline
6. Locus of Control (Rotter and WVS) – we predict less IPV in a relationship and attenuated treatment effects in households in which the female and/or the male respondent exhibit internal locus of control at baseline
7. Coping Strategies (Folkman & Lazarus) – we predict less IPV in a relationship and attenuated treatment effects in households in which the female and/or the male respondent exhibit active coping strategies at baseline
8. Childhood history of violence – we predict greater IPV in a relationship and attenuated treatment effects for households in which the female and/or the male respondent experienced violence during childhood

4.7.2 Mediation and moderation analyses for IPV outcomes

To assess possible pathways through which the transfers affected rates of physical and sexual IPV, quantitative and qualitative exploratory analyses will be performed. Semi-structured

interviews and focus groups were conducted to examine changes in the household and in intimate partner dynamics following the introduction of the cash transfers in treatment and spillover households. Interview guide topics include household decision-making, relationship quality, marital conflict and peace, perceived IPV norms, and changes experienced at the household and community levels following cash transfer disbursement. Transcripts will be analyzed using thematic analysis and text analysis to examine possible mechanisms of change in IPV. Quantitatively, mediation analyses of treatment effects on female reports of experienced physical and sexual violence may be conducted, as a hypothesis-generating exercise, with a range of psychological, social, interpersonal, and financial variables. Examples include the following variables: psychosocial resources (e.g. depressive symptoms, optimism, happiness, self-esteem, perceived stress, self-efficacy, coping strategies); social capital; female bargaining power; female economic empowerment; joint decision-making; marital relationship satisfaction; emotional violence; marital control; respect by partner, family, and community; community responses to cash transfers; financial strain; food security; alcohol use; sources of conflict; adherence to gender roles; male-focused attitudes; justifiability of violence; power perceptions; and perceived gender norms. Moderating variables that will be considered include: childhood history of violence; stressful life events; duration of cohabitation; inception of violence in relationship; marital status; perceived male partner aggressiveness; age of partners; age disparity between partners; education level of partners; education disparity level between partners; household size and composition; psychological variables at baseline (e.g. depressive symptoms, optimism, happiness, self-esteem, perceived stress); and domestic violence and bargaining power at baseline. Predictive models may include such variables as well as variables that may be considered proxies of violence against the female, such as reports of physical injury, list method reports of violence, and reports of village-level rate of IPV.

4.8 Attrition

To assess whether attrition potentially confounds our results, we proceed as follows. First, we define the variable $attrit_{vhQ_1Q_2}$ for households that attrited between any two survey rounds Q_1 and Q_2 . For treatment and spillover households, Q_1 may be either baseline or endline 1, Q_2 may be either endline 1 or endline 2. For pure control households, Q_1 is endline 1 and Q_2 is endline 2. For the results to be unaffected by attrition, we require (i) attrition is minimal and not differential across treatment groups between endline 1 and endline 2 for treatment, spillover and pure control households, and (ii) attrition is minimal and not differential across treatment groups between baseline and endline 2 for treatment and spillover.

We then assess the severity of attrition using three approaches. The following equations estimate whether the magnitude of attrition is different for treatment, spillover and pure control households across the relevant survey rounds:

$$attrit_{vhBE_1} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhE_1} \quad (21)$$

$$attrit_{vhE_1E_2} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \beta_2 S_{vh} + \varepsilon_{vhE_2} \quad (22)$$

$$attrit_{vhBE_2} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhE_2} \quad (23)$$

Second, equation 24 assesses whether attrition households are different in terms of a comprehensive range of baseline characteristics:

$$y_{vhB} = \alpha_v + \beta_0 + \beta_1 attrit_{vhBE_1} + \varepsilon_{vhB} \quad (24)$$

$$y_{vhB} = \alpha_v + \beta_0 + \beta_1 attrit_{vhBE_2} + \varepsilon_{vhB} \quad (25)$$

$$y_{vhE_1} = \alpha_v + \beta_0 + \beta_1 attrit_{vhE_1E_2} + \varepsilon_{vhE_1} \quad (26)$$

And third, equation 27 measures whether the baseline characteristics of attrition households are significantly different across treatment groups. The sample for regression will be restricted to attrition households:

$$(y_{vhB} \mid attrit_{vhBE_1} = 1) = \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhB} \quad (27)$$

$$(y_{vhB} \mid attrit_{vhBE_2} = 1) = \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhB} \quad (28)$$

$$(y_{vhE_1} \mid attrit_{vhE_1E_2} = 1) = \beta_0 + \beta_1 T_{vh} + \beta_2 S_{vh} + \varepsilon_{vhE_1} \quad (29)$$

Where the outcomes are at the individual level, standard errors will be clustered at the household level, with appropriate weighting. If worrying levels of attrition are found, we will adjust for the potential effect of such attrition using Lee bounds.

4.9 Accounting for multiple inference

As cash transfers are likely to impact a large number of economic behaviors and dimensions of welfare, and given that our survey instrument often included several questions related to a single behavior or dimension, we will account for multiple hypotheses by using outcome variable indices and family-wise p -value adjustment.

We have catalogued below the primary groups of outcomes that we intend to consider in the analysis. For each of these outcome groups, we will construct either an index variable following the procedure proposed by Anderson (2008), or choose a focal variable. For these indices and focal variables, we will report both unadjusted p -values, and p -values adjusted for multiple comparisons using Anderson’s (2008) variant of Efron & Tibshirani’s (1993) non-parametric permutation test. In addition, to make our analyses comparable to the anticipated analyses in the new study by Haushofer, Miguel, Niehaus, and Walker, we will create “naïve” indices by simply averaging outcomes, and will report p -values adjusted for the false discovery rate. Within groups of outcomes, we report unadjusted p -values.

5 Outcome variables

5.1 Components of indices and focal variables

The following indices and focal variable will be the primary focus of our analysis and will be included in multiple comparison adjustments.

Total assets: Total value in 2012 PPP adjusted dollars of all household assets:

1. Moveable assets
 - (a) Livestock
 - i. Cows
 - ii. Small livestock
 - iii. Birds
 - (b) Furniture
 - (c) Agricultural tools
 - (d) Radio or TV
 - (e) Other assets
2. Savings
3. Value of roof (inclusion not pre-specified for endline 1 analysis)
4. Omitted: Value of land (omission not pre-specified for endline 1 analysis)

Total consumption: Total spending per month in 2012 PPP adjusted dollars:

1. Food
 - (a) Food own production
 - (b) Food bought
 - (c) Meat & fish
 - (d) Fruit & vegetables
 - (e) Other food
2. Temptation good expenditure
3. Medical expenditure
 - (a) Medical expenditure (respondent)
 - (b) Medical expenditure (spouse)
 - (c) Medical expenditure (children)
4. Education expenditure
5. Social expenditure
6. Omitted: Durables expenditure, house expenditure (omission not pre-specified for end-line 1 analysis)
7. Other expenditure

Agricultural and business income: Total household enterprise revenue per month in 2012 PPP adjusted dollars:

1. Agricultural income
 - (a) Agricultural income (own consumption, total)
 - i. Agricultural income (own consumption, harvest)
 - ii. Agricultural income (own consumption, animals)
 - (b) Agricultural income (sales, total)
 - i. Agricultural income (sales, harvest)
 - ii. Agricultural income (sales, animal products)
 - iii. Agricultural income (sales, animals)
2. Non-farm enterprise revenue

Psychological wellbeing index: Standardized weighted average of psychological and neurobiological measures:

1. Depression (CESD) - negatively coded
2. Worries - negatively coded
3. Stress (Cohen) - negatively coded
4. Happiness (WVS)
5. Life satisfaction (WVS)
6. Cortisol (in log nm/l adjusted for confounds) - negatively coded

Food security index (household): Weighted average of measures of food security and hunger:

1. Meals skipped in the last month (adults) - negatively coded
2. Whole days without food in the last month (adults) - negatively coded
3. Meals skipped in the last month (children) - negatively coded
4. Whole days without food in the last month (children) - negatively coded
5. Household ate less preferred/cheaper foods in the last month (# of times) - negatively coded
6. Household relied on help from others for food in the last month (# of times) - negatively coded
7. Household purchased food on credit in the last month (# of times) - negatively coded
8. Household had to hunt, gather wild food, harvest prematurely in the last month (# of times) - negatively coded
9. Household begged because not enough food in the house in the last month (# of times) - negatively coded
10. All members usually eat two meals (dummy)
11. All members usually eat until content (dummy)
12. Number of times ate meat or fish (last week)

13. Enough food in the house for tomorrow (dummy)
14. Respondent slept hungry in the last week (dummy) - negatively coded
15. Respondent ate protein in the last 24 hours (dummy)
16. Proportion of HH who ate protein in the last 24 hours
17. Proportion of children who ate protein in the last 24 hours

Health index: Standardized weighted average:

1. Proportion of household sick/injured - negatively coded
2. Proportion of children sick/injured - negatively coded
3. Proportion of sick/injured who could afford treatment
4. Proportion of illnesses where doctor was consulted
5. Proportion of newborns vaccinated
6. Proportion of children < 14 getting checkup in the last 6 months
7. Proportion of children < 5 who died in the past 12 months - negatively coded

Education index: Standardized weighted average:

1. Education expenditure per child
2. Proportion of school-aged children in school

Female empowerment index: Standardized weighted average of attitude index and violence index:

1. Violence index (standardized weighted average):
 - (a) Female report of number of instances of physical violence - negatively coded
 - (b) Female report of number of instances of sexual violence - negatively coded
 - (c) Female report of number of instances of emotional violence - negatively coded
2. Attitudes index (standardized weighted average):
 - (a) Justifiability of violence score - negatively coded
 - (b) Male-focused attitudes score - negatively coded

5.2 All variables collected

In addition to the indices which are our primary focus, we will also explore impacts on other variables collected. Analysis of the variables not included in the above indexes is intended to be exploratory. We will look at these variables to better understand mechanisms and to generate hypotheses for future work.

1. Assets

(a) Movable assets

i. Livestock: Sum of all livestock assets owned by respondents in KES (later converted to USD PPP), including cows, small livestock, and birds.

ii. Furniture: Value of cupboards, sofas, chairs, tables, clocks, stoves, and beds as self reported in KES (later converted to USD PPP).

iii. Agricultural tools: Value of farming tools, wheelbarrows, and hand carts, in KES (later converted to USD PPP).

iv. Radio or TV: Value of radio and television assets in KES (later converted to USD PPP)

v. Other assets: Value of bicycles, motorbikes, solar panels, cellphones, and any other assets that respondents reported when asked if they owned any additional assets apart from those listed, in KES (later converted to USD PPP).

(b) Savings: Value of savings, in KES (later converted to USD PPP), in all savings accounts for the household (including mobile money accounts).

(c) Land owned: Land owned in acres.

(d) House has non-thatch roof: Dummy variable indicating that responding has a non-thatch roof (i.e. iron sheets, wood, etc.)

(e) House has non-mud floor: Dummy variable indicating that respondent has floor consisting of materials other than mud (i.e. tiles, wood, stones, concrete, etc.)

(f) **House has non-mud walls:** Dummy variable indicating that respondent has wall constructed from materials other than mud (i.e. wood, bricks/stones, plaster/cement).

(g) **House has electricity:** Dummy variable indicating that respondent has electricity

(h) **House has toilet or pit latrine:** Dummy variable indicating that the respondent has a pit latrine or mobile / portable toilet.

2. Consumption

(a) Food

i. Food own production: Value of milk consumed, other animal products consumed (cattle, small livestock, birds), meat consumed (cattle, small livestock, birds), eggs consumed, as well as the value of the crops consumed both for the long rains and short rains seasons, on average per week in KES (later converted to USD PPP).

ii. Food bought: Value of cereals, vegetables, fruit, meat, fish, dairy, fats, sugars, drinks, spices, and prep food purchased in the past week in KES (later converted to USD PPP).

ii.a. Meat & fish: Value of meat and fish purchased in the past week in KES (later converted to USD PPP).

ii.b. Fruit & vegetables: Value of fruits and vegetables purchased in the past week in KES (later converted to USD PPP).

ii.c. Other food: Value of cereals, dairy products, fats, prep foods, drinks, and spices purchased in the past week in KES (later converted to USD PPP).

(b) **Temptation good expenditure:** Value of expenditure on alcohol, tobacco, and lottery tickets in the past week in KES (later converted to USD PPP).

(c) **Medical expenditure:** Value of medical expenditure (consultation fees, medicines, hospitalizations) for the respondent, spouse, and children of the respondent in the past 1 month, in KES (later converted to USD PPP).

i. Medical expenditure (respondent): Value of medical expenditures (consultation fees, medicines, hospitalizations) in the past 1 month in KES (later converted to USD PPP) for the respondent.

ii. Medical expenditure (spouse): Value of medical expenditures (consultation fees, medicines, hospitalizations) in the past 1 month in KES (later converted to USD PPP) for the spouse of the respondent.

iii. Medical expenditure (children): Value of medical expenditures (consultation fees, medicines, hospitalizations) in the past 1 month in KES (later converted to USD PPP) for the children of the respondent.

(d) Education expenditure: Value of education costs consumed (school fees, uniforms, etc.) in the past 12 months in KES (later converted to USD PPP).

(e) Durables expenditure: Value of household durables (cutlery, pots/pans, light bulbs, curtains, carpets, etc.) in the past 12 months in KES (later converted to USD PPP).

(f) House expenditure: Value of expenditure on house/land rent and repair in the past 12 months in KES (later converted to USD PPP).

(g) Social expenditure: Value of expenditure on ceremonies, weddings, funerals, dowry, village elders, and any other recreation (cinema tickets, music/CDs, books/magazines, etc.) in the past 12 months in KES (later converted to USD PPP).

(h) Other expenditure: Value of expenditure on airtime, traveling (petrol, bus fare, hotel stays), clothing, personal items (haircut, hair oil, cosmetics, etc.), household items (soap, toilet paper, candles, etc.), firewood, electricity bill, and water bills in the past 1 month in KES (later converted to USD PPP).

3. Food security

(a) Meals skipped (adults): Frequency of adults having to cut the size of meals or skip them entirely in the past 1 month.

(b) Whole days without food (adults): Frequency that adults have gone without any meals by in the past month.

(c) Meals skipped (children): Frequency of children (<14 years of age) having to cut the size of meals or skip them entirely in the past 1 month.

(d) Whole days without food (children): Frequency that children (<14 years of age) have gone without any meals by in the past month.

(e) Eat less preferred / cheaper foods: Frequency that household members have had to eat less preferred or less expensive foods in the past month.

(f) Rely on help from others for food: Frequency that household members have had to borrow food or rely on help from a friend or relative in the past month.

(g) Purchase food on credit: Frequency that household members have had to purchase food on credit.

(h) Hunt, gather wild food, harvest prematurely: Frequency that household members have had to gather wild food, hunt, or harvest immature crops in the past month.

(i) Beg because not enough food in the house: Frequency of household members having to beg because there was not enough food in the household in the past month.

(j) All members eat two meals: Dummy variable indicating whether all members of the household regularly eat at least 2 meals a day.

(k) All members eat until content: Dummy variable indicating whether all members usually eat until they are content each day.

(l) Number of times ate meat or fish: Frequency of respondent eating meat, eggs, or fish in the last week.

(m) Enough food in the house for tomorrow?: Dummy variable indicating whether the respondent believes that the household has enough food for tomorrow.

(n) Respondent slept hungry: Dummy variable indicating whether the respondent has gone to sleep hungry in the past week.

(o) Respondent ate protein: Dummy variable indicating whether the respondent ate protein in the past week.

(p) Proportion of household who ate protein: Number of people listed by respondent as having eaten protein in the past week divided by the total number of members in the household.

(q) Proportion of children who ate protein: Number of children listed by respondent (including own children and stepchildren) who ate protein divided by the total number of children in the household.

4. Psychological and neurobiological outcomes

(a) Depression (CES-D)

(b) Worries

(c) Stress (Cohen)

(d) Happiness (WVS)

(e) Life satisfaction (WVS)

(f) Cortisol

(g) Trust (WVS)

(h) Locus of control (Rotter and WVS)

(i) Optimism (Scheier)

(j) Self-esteem (Beck)

(k) Self-Efficacy (Schwarzer & Jerusalem)

(l) Sense of power (Anderson, John, & Keltner)

(i) **Aspirations** Standardized weighted average of the following:

1. Income aspired to and likely to achieve
2. Assets aspired to and likely to achieve
3. Social status aspired to and likely to achieve
4. Education (oldest child) aspired to and likely to achieve

5. Female empowerment

(a) **Physical violence dummy:** Dummy indicating if any physical violence occurred in the last six months, including if the spouse pushed, twisted the arm of, punched, kicked, chokes, or pulled a knife on the respondent.

(b) **Sexual violence dummy:** Dummy indicating if any sexual violence occurred in the last six months, including if the spouse raped or performed non-consensual sexual acts on the respondent.

(c) **Emotional violence dummy:** Dummy indicating if any emotional violence occurred in the last six months, including if the spouse was jealous or angry if you talked to other men/women, accused you of being unfaithful, did not permit you to meet your friends of the same gender, tried to limit your contact with your family, or did not trust you with any money.

(d) **Justifiability of violence score:** Total number of situations in which the respondent feels that the husband is justified in beating his wife: can beat if he/she goes out without telling her, if he/she neglects the children, he/she argues with her, he/she refused to have sex with him/her, he/she burns the food. Additional scenarios included in the follow-up.

(e) **Male-focused attitudes score:** Sum of all dummy variables indicating whether the respondent agree with the following male oriented statements: men should make the important decisions in the family, the wife has the right to express her opinion even when she disagrees with her husband (reverse coded), wife should tolerate getting beaten to keep family together, husband has the right to beat his wife, it is more important to send a son to school than to send a daughter. Additional scenarios included in the follow-up.

(f) Male makes decisions: Sum of dummy situations in which the respondent believes the male should have the final say: contraception use, children's schooling, buying clothes or shoes, what to do if a child falls ill, disciplining children, whether to have children, how much to spend on food, extra spending, saving.

(g) Proportion choosing money for spouse vs. self: Number of respondents choosing to give their spouse 130 KES vs. keeping 100 KES (later converted to USD PPP) for themselves divided by total number of married respondents.

(h) Physical violence frequency: Number of incidents of physical violence in the last six months, including if the spouse pushed, twisted the arm of, punched, kicked, chokes, or pulled a knife on the respondent in the past six months.

(i) Sexual violence frequency: Number of incidents of sexual violence in the last six months, including if the spouse raped or performed non-consensual sexual acts on the respondent in the past six months.

(j) Emotional violence frequency: Number of incidents of emotional violence in the last six months, including if the spouse was jealous or angry if you talked to other men/women, accused you of being unfaithful, did not permit you to meet your friends of the same gender, tried to limit your contact with your family, or did not trust you with any money.

(k) Economic control frequency: Number of occurrences of the following: including whether husband expected you to ask permission to purchase large or small items, took your earnings against your will, tells you he does not have enough money to give you for household expenses, tells you he does not have enough money to give you to spend on yourself, refuses to give you money for household expenses, even when he had money for other things, require that you give up or refuse a job for money outside the home because he did not want you to work, make important financial decisions without talking to you about them, demand to know how you spent money, hide money from you, spent money set aside for household benefits on himself, threaten not to give you money or take it away from you, given you little money or reduced your spacing when he is angry.

(l) Perceived village IPV dummy Dummy indicating an affirmative response to the question "Do men in your village beat, slap, or act physically violent towards their wives?"

(j) Perceived village IPV frequency Numerical response to the question “How many times per month do you think a man in your village beats, slaps or acts violently towards his wife?”

(k) Physical injury dummy A dummy variable indicating whether individual answers yes to whether the respondent experienced any of the following as a result of an act by her partner: had cuts or bruises; had eye injuries, sprains, dislocations, or burns; had deep wounds, broken bones, broken teeth or any other serious injury.

(l) Perceived community justifiability of violence score Total number of situations in which the respondent feels that most individuals in her community believe the husband is justified in beating his wife: can beat if he/she goes out without telling her, if he/she neglects the children, he/she argues with her, he/she refused to have sex with him/her, he/she burns the food. Additional scenarios included in the follow-up.

(m) Perceived community male-focused attitudes score Sum of all dummy variables indicating whether the respondent believes most individuals in her community agrees with the following male oriented statements: men should make the important decisions in the family, the wife has the right to express her opinion even when she disagrees with her husband (reverse coded), wife should tolerate getting beaten to keep family together, husband has the right to beat his wife, it is more important to send a son to school than to send a daughter. Additional statements included in the follow-up.

(n) Marital control Frequency with which the respondent has experience one of the following in the last six months: was jealous if you talked to other men, accused your of being unfaithful, did not permit you to meet female friends, tried to limit your contact with your family, did not trust you with money, insisted on knowing where you were at all times, expected you to ask permission before leaving the house.

(o) Who in household owns livestock assets Dummy variable for whether livestock are owned by the wife or jointly by the husband and wife.

(p) Who manages money from livestock income Dummy variable for whether livestock income is managed by the wife or jointly by the husband and wife.

(q) Who controls savings accounts Dummy variable for whether savings accounts are managed by the wife or jointly by the husband and wife.

(r) **Who borrows** Dummy variable for whether borrowing decisions are made by the wife or jointly by the husband and wife.

6. Health

(a) **Medical expenses per episode (entire household):** Sum of all treatment costs (direct and indirect) in KES (later converted to USD PPP) for any episodes in the past month among all household members divided by the total number of incidents in the household.

(b) **Medical expenses per episode (spouse):** Sum of all treatment costs (direct and indirect) in KES (later converted to USD PPP) for any episodes in the past month among spouses in the household divided by the total number of incidents among spouses in the household.

(c) **Medical expenses per episode (children):** Sum of all treatment costs (direct and indirect) in KES (later converted to USD PPP) for any episodes in the past month among spouses in the household divided by the total number of incidents among children in the household.

(d) **Proportion of household sick / injured:** Total number of household members who were sick or injured in the past month divided by the total number of household members.

(e) **Proportion of children sick / injured:** Total number of children in the household who were sick or injured in the past month divided by the total number of children in the household.

(f) **Proportion of sick / injured who could afford treatment:** Total number of household members who were sick / injured who reported being able to pay for treatments divided by the total number of people who reported being sick/injured in the past month.

(g) **Average number of sick days per household member:** Total number of sick days among household members divided by the number of household members in the past month.

(h) **Proportion of illnesses where doctor was consulted:** Total number of illness/injury episodes where a doctor was consulted divided by the total number of illnesses and injuries in the household in the past month.

(i) Proportion of newborns vaccinated: Total number of children under one years of age who have been vaccinated divided by the total number of children under one years of age in the household.

(j) Proportion of children <14 getting checkup: Total number of children under the age of 14 reporting having a regular checkup in the past six months divided by the total number of children under the age of 14.

(k) Proportion of children <5 who died: Total number of children in the household who have died in the past twelve months divided by the total number of children under 5 (living and passed) in the household.

7. Education

(a) Total education expenditure: Value spend on educations goods (school fees, uniforms, books, or other supplies, in KES (later converted to USD PPP) for the household in the past 12 months.

(b) Education expenditure per child: Value spent on education goods (school fees, uniforms, books, or other supplies, in KES (later converted to USD PPP) for the household in the past 12 months divided by the number of school age children (aged 3-18) in the household.

(c) Proportion of school-aged children in school: Number of school age children (aged 3-18) currently attending school divided by the total number of school age children in the household.

(d) School days missed for economic reasons, per child: Sum of total number of days per child reported as missed for economic reasons (No breakfast / food, can't pay fees, needs to work for money, needed for household, child or elder care) divided by the total number of school aged children in the past month.

(e) Income generating activities per school-aged child >6: Sum of total number of income generating activities per child 6-18 years of age in the household divided by the number of children 6-18 in the household engaged in the past twelve months.

8. Enterprise

(a) Agricultural income (total)

i. Agricultural income (own consumption, total): Sum of consumed harvest income and consumed animal income in KES (later converted to USD PPP) per month.

ii. Agricultural income (sales, total): Sum of harvest sales, animal product sales, and livestock sales to create a monthly agricultural income average.

(b) Enterprise profits (6 months): Value in KES (later converted to USD PPP) of profits (or losses if negative) of all non-agricultural, non-livestock income generating enterprises owned and operated (partially or fully) by the respondent in the past six months.

(c) Enterprise revenue (1 month): Value in KES (later converted to USD PPP) of all money received from all non-agricultural, non-livestock income generating enterprises owned and operated (partially or fully) by the respondent in the past one month.

(d) Enterprise revenue (typical month): Value in KES (later converted to USD PPP) of the sales of all non-agricultural, non-livestock income generating enterprises owned and operated (partially or fully) by the respondent in an average month.

(e) New non-agricultural business owner (dummy): Dummy variable indicating whether a respondent did not have a non-agricultural business at baseline but now does at endline.

(f) Non-agricultural business owner (dummy): Dummy variable indicating whether a respondent owns and operates a non-agricultural business.

(g) Number of employees: Number of non-household member employees in all entrepreneurial activities owned and operated by the respondent (partially or fully owned).

(h) Value of investment in non-agricultural income (total): Costs of electricity, wages, water, transport ,inputs, and any other expenses for all enterprises owned and operated (partially or fully) by the respondent for the past three months in KES (later converted to USD PPP).

9. Financial variables

(a) Value of outstanding loans: Amount in KES (later converted to USD PPP) outstanding from any loan taken by a member of the household, including debts to local shops and kiosks.

(b) Unable to pay loans (12 months): Dummy variable indicating that household was unable to make payments on at least one loan in the past 12 months

(c) Value of remittance sent: Value of all cash and goods sent as remittances to non-household members or members outside of their compound in the past month in KES (later converted to USD PPP).

(d) Value of remittances received: Value of all cash and goods received as remittances from non-household members or members outside of their compound in the past month in KES (later converted to USD PPP).

(e) Net remittances: Value of remittances sent less value of remittances received in KES (later converted to USD PPP).

10. List method

(a) Temptation goods: Estimated number of alcohol and tobacco users in treatment and control groups.

(b) Intimate partner violence: Estimated rate of violence in treatment and control groups.

11. Political Variables

(a) Will vote in the next election: Indicator for answering yes to “will you be voting in the upcoming national elections that will be held next year?”

(b) Political knowledge: Indicator for knowing the names of the candidates running for Prime Minister and President in the next election.

(c) Attitudes towards voting: Indicator for responding that it is very Kenyan citizen’s responsibility to vote when asked about responsibility to vote.

(d) Trust in government institutions: Indicator for answering “let the Kenyan government decide how to spend it” when asked the how foreign aid should be spent to reduce poverty.

12. Labor and Time Use Variables:

(a) Salaried jobs:

i. Salaried labor is the household’s primary source of income: Indicator for answering that a salaried job is household’s primary source of income.

ii. Proportion of household members working in a salaried job: Proportion of adults in the household for who were reported as having worked in a salaried job at any point in the last 12 months.

iii. Time (in days) spent working in a salaried job Days in the last month spent working a salaried job by household adults.

iv. Income working in a salaried job Typical income the last month from working a salaried job by household adults.

(b) Casual Labor for other households

Time spent by household adults performing casual labor, performing housework for pay, farming land, or tending animals for another household.

i. Casual labor is the household’s primary source of income: Indicator for answering that casual labor for other households is household’s primary source of income.

ii. Proportion of household members performing casual labor: Proportion of adults in the household for who were reported as having performed casual labor at any point in the last 12 months.

iii. Time (in days) spent performing casual labor Days in the last month spent performing casual labor by household adults.

iv. Income from performing casual labor Typical income the last month from performing casual labor by household adults.

(c) Household enterprises:

Time spent by household adults working in a business owned by a household member, farming on land owned by the household, tending animals owned by the household, performing housework in this household, or fishing.

i. Household or agricultural enterprise is the household's primary source of income: Indicator for answering that a household / agricultural enterprise is household's primary source of income.

ii. Proportion of household members working for household or agricultural enterprise: Proportion of adults in the household for who were reported as having worked in household / agricultural enterprise at any point in the last 12 months.

iii. Time (in days) spent working for household or agricultural enterprise Days in the last month spent working in household / agricultural enterprise by household adults.

iv. Income from household / agricultural enterprise Typical income the last month from working in household / agricultural enterprise.

(d) Time use

i. Time working for HH in days Farming fishing, tending animals, housework, working for an enterprise owned by the household

ii. Housework in days Performing housework for pay for another household

iii. Time working outside the HH in days Farming fishing, tending animals, housework, daily labor, salaried job

iv. Income from work outside the HH in days Farming fishing, tending animals, housework, daily labor, salaried job

13. Social Capital

(a) Group membership Total number of types of groups in which the individual has been active in the last 12 months: work related / trade union, community associations, women's groups or Chama, political groups, religious groups, credit /funeral groups, sports groups, other.

(b) Support from groups Total number of types of groups from which the individual has received emotional help, economic help, or assistance in learning to do things: work related / trade union, community associations, women's groups or Chama, political groups, religious groups, credit /funeral groups, sports groups, other.

(c) Support from individuals Total number of individuals from which the individual has received emotional help, economic help, or assistance in learning to do things: family, neighbors, friends, community leaders, religious leaders, politicians, government officials / civil servants, charities / NGOs.

(d) Collective action Indicator for whether the individual had joined together with other community members to address a problem or common issue.

(e) Trust Indicator for whether the individual believes the majority of people in the community can be trusted.

14. Impact of transfers on relationships

(a) Intimate partner relationships Relationship satisfaction, respect, jealousy, marital status, quality of relationship, conflict; decision-making on use of transfer; use of transfers

(b) Family relationships Respect from in-laws

(c) Community relationships Tension between household and community; respect from other households; change in how household was treated; change in relationship due to the transfer

(d) Marital satisfaction Total score across marital satisfaction questions

(e) Social comparisons

References

- Aguëro, J. M., Michael Carter, and I. Woolard. 2010. “The impact of unconditional cash transfers on nutrition: The South African Child Support Grant.” Southern Africa Labour and Development Research Unit Working Paper 06/08.
- Aker, Jenny C. 2015. “Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo.” *The World Bank Economic Review*, November, lhv055.
- Aker, Jenny C., Rachid Boumnijel, Amanda McClelland, and Niall Tierney. 2016. “Payment Mechanisms and Anti-Poverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger.” *Economic Development and Cultural Change (forthcoming)*.
- Akresh, Richard, Damien de Walque, and Harounan Kazianga. 2013. “Cash transfers and child schooling: Evidence from a randomized evaluation of the role of conditionality.” Policy Research Working Paper 6340, The World Bank.
- Amarante, Veronica, Rodrigo Arim, Gioia De Melo, and Andrea Vigorito. 2010. “Family allowances and child school attendance: An ex-ante evaluation of alternative schemes in Uruguay.” In *Child Welfare in Developing Countries*, edited by John Cockburn and Jane Kabubo-Mariara, 211–245. New York: Springer.
- Anderson, Michael. 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.
- Ardington, Cally, Anne Case, and Victoria Hosegood. 2009. “Labor supply responses to large social transfers: Longitudinal evidence from South Africa.” *American Economic Journal: Applied Economics* 1 (1): 22–48.
- Attanasio, Orazio, and Alice Mesnard. 2006. “The impact of a conditional cash transfer programme on consumption in Colombia.” *Fiscal Studies* 27 (4): 421–442.
- Baird, Sarah, Ephraim Chirwa, Jacobus De Hoop, and Berk Özler. 2016. “Girl power: Cash transfers and adolescent welfare. Evidence from a cluster-randomized experiment in Malawi.” In *African Successes: Human Capital, Volume 2*. University of Chicago Press.
- Baird, Sarah, Jacobus De Hoop, and Berk Özler. 2013. “Income shocks and adolescent mental health.” *Journal of Human Resources* 48 (2): 370–403.
- Baird, Sarah, Craig McIntosh, and Berk Özler. 2011. “Cash or condition? Evidence from a cash transfer experiment.” *The Quarterly Journal of Economics* 126 (4): 1709–1753.
- Baird, Sarah, and Berk Özler. 2012. “Examining the reliability of self-reported data on school participation.” *Journal of Development Economics* 98 (1): 89–93 (May).
- Baird, Sarah J., Richard S. Garfein, Craig T. McIntosh, and Berk Özler. 2012. “Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: A cluster randomised trial.” *The Lancet* 379 (9823): 1320–1329.

- Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2013. "Can Basic Entrepreneurship Transform the Economic Lives of the Poor?" SSRN Scholarly Paper ID 2266813, Social Science Research Network, Rochester, NY.
- Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Pariente, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A multi-faceted program causes lasting progress for the very poor: Evidence from six countries." *Science* 348, no. 6236 (May).
- Barham, Tania, Karen Macours, and John A. Maluccio. 2013. "More schooling and more learning?: Effects of a three-year conditional cash transfer program in Nicaragua after 10 years." IDB Working Paper Series IDB-WP-432, Inter-American Development Bank.
- Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating skilled employment in developing countries: Experimental evidence from Uganda." *Quarterly Journal of Economics* 129 (2): 697–752.
- Blattman, Christopher, Julian Jamison, and Margaret Sheridan. 2015. "Reducing crime and violence: Experimental evidence on adult noncognitive investments in Liberia." NBER Working Paper 21204.
- Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. 2015. "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change* 64 (2): 187–220 (December).
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust inference with multiway clustering." *Journal of Business & Economic Statistics* 29 (2): 238–249.
- Cunha, Jesse M. 2014. "Testing paternalism: Cash versus in-kind transfers." *American Economic Journal: Applied Economics* 6 (2): 195–230.
- Currie, Janet, and Firouz Gahvari. 2008. "Transfers in cash and in kind: Theory meets the data." *Journal of Economic Literature* 46 (2): 333–383.
- De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to capital in microenterprises: Evidence from a field experiment." *The Quarterly Journal of Economics* 123 (4): 1329–1372.
- Duflo, Esther. 2003. "Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa." *The World Bank Economic Review* 17 (1): 1–25.
- Edmonds, Eric V. 2006. "Child labor and schooling responses to anticipated income in South Africa." *Journal of Development Economics* 81 (2): 386–414.
- Edmonds, Eric V., and Norbert Schady. 2012. "Poverty alleviation and child labor." *American Economic Journal: Economic Policy* 4 (4): 100–124.
- Efron, Bradley, and Robert Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press.
- Fafchamps, Marcel, David McKenzie, Simon R. Quinn, and Christopher Woodruff. 2011. "When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana." NBER Working Paper 17207.

- Fafchamps, Marcel, and Simon Quinn. 2015. "Aspire." NBER Working Paper 21084.
- Fernald, Lia CH, and Melissa Hidrobo. 2011. "Effect of Ecuador's cash transfer program Bono de Desarrollo Humano on child development in infants and toddlers: A randomized effectiveness trial." *Social Science & Medicine* 72 (9): 1437–1446.
- Ferreira, Francisco H. G., Deon Filmer, and Norbert Schady. 2009. "Own and sibling effects of conditional cash transfer programs: Theory and evidence from Cambodia." Policy Research Working Paper 5001, The World Bank.
- Filmer, Deon, and Norbert Schady. 2009. "Are There Diminishing Returns to Transfer Size in Conditional Cash Transfers?" Policy Research Working Paper 4999, The World Bank.
- Gertler, Paul, Sebastian Martinez, and Marta Rubio-Codina. 2012. "Investing cash transfers to raise long-term living standards." *American Economic Journal: Applied Economics* 4 (1): 164–192.
- Jack, William, and Tavneet Suri. 2014. "Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution." *The American Economic Review* 104 (1): 183–223.
- Macours, Karen, Norbert Schady, and Renos Vakis. 2012. "Cash transfers, behavioral changes, and cognitive development in early childhood: Evidence from a randomized experiment." *American Economic Journal: Applied Economics* 4 (2): 247–73.
- Maluccio, John. 2010. "The impact of conditional cash transfers on consumption and investment in Nicaragua." *Journal of Development Studies* 46 (1): 14–38.
- Maluccio, John, and Rafael Flores. 2005. "Impact evaluation of a conditional cash transfer program: the Nicaraguan Red de Proteccion Social." Research Report 141, International Food Policy Research Institute (IFPRI).
- Maluccio, John, Alexis Murphy, and Ferdinando Regalia. 2010. "Does supply matter? Initial schooling conditions and the effectiveness of conditional cash transfers for grade progression in Nicaragua." *Journal of Development Effectiveness* 2 (1): 87–116.
- McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99 (2): 210–221.
- Mckenzie, David J. 2015. "Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition." Policy Research Working Paper 7391, The World Bank.
- Oosterbeek, Hessel, Juan Ponce, and Norbert Schady. 2008. "The impact of cash transfers on school enrollment: Evidence from Ecuador." Policy Research Working Paper 4645, The World Bank.
- Paxson, Christina, and Norbert Schady. 2010. "Does money matter? The effects of cash transfers on child development in rural Ecuador." *Economic Development and Cultural Change* 59 (1): 187–229.
- Pepper, John V. 2002. "Robust inferences from random clustered samples: An application using data from the panel study of income dynamics." *Economics Letters* 75 (3): 341–345.

- Posel, Dorrit, James A. Fairburn, and Frances Lund. 2006. "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa." *Economic Modelling* 23 (5): 836–853.
- Sadoulet, Elisabeth, Alain de Janvry, and Benjamin Davis. 2001. "Cash transfer programs with income multipliers: PROCAMPO in Mexico." *World Development* 29 (6): 1043–1056.
- Schady, Norbert, and Maria Caridad Araujo. 2008. "Cash transfers, conditions, and school enrollment in Ecuador." *Economía* 8 (2): 43–70.
- Schady, Norbert, and José Rosero. 2008. "Are cash transfers made to women spent like other sources of income?" *Economics Letters* 101 (3): 246–248.
- Skoufias, Emmanuel, Mishel Unar, and Teresa Gonzalez de Cossio. 2013. "The poverty impacts of cash and in-kind transfers: experimental evidence from rural Mexico." *Journal of Development Effectiveness* 5 (4): 401–429 (December).
- Team, The Kenya CT-OVC Evaluation. 2012a. "The impact of Kenya's Cash Transfer for Orphans and Vulnerable Children on human capital." *Journal of Development Effectiveness* 4 (1): 38–49.
- . 2012b. "The impact of the Kenya Kenya's Cash Transfer for Orphans and Vulnerable Children on household spending." *Journal of Development Effectiveness* 4 (1): 9–37.
- Zwane, Alix, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean Karlan, Richard Hornbeck, Xavier Giná, Esther Duflo, Florencia Devoto, Bruno Crepon, and . 2011. "Being surveyed can change later behavior and related parameter estimates." *Proceedings of the National Academy of Sciences* 108 (5): 1821–1826.