THE SHORT-TERM IMPACT OF UNCONDITIONAL CASH TRANSFERS TO THE POOR: EXPERIMENTAL EVIDENCE FROM KENYA

Johannes Haushofer, **Jeremy Shapiro**

Princeton University and Busara Center for Behavioral Economics, Busara Center for Behavioral Economics

Abstract

We use a randomized controlled trial to study the response of poor households in rural Kenya to unconditional cash transfers from the NGO GiveDirectly. The transfers differ from other programs in that they are explicitly unconditional, large, and concentrated in time. We randomized at both the village and household levels; furthermore, within the treatment group, we randomized recipient gender (wife versus husband), transfer timing (lump-sum transfer versus monthly installments), and transfer magnitude (US$404 PPP versus US$1,525 PPP). We find a strong consumption response to transfers, with an increase in household monthly consumption from $158 PPP to $193 PPP nine months after the transfer began. Transfer recipients experience large increases in psychological well-being. We find no overall effect on levels of the stress hormone cortisol, although there are differences across some subgroups. Monthly transfers are more likely than lump-sum transfers to improve food security, whereas lump-sum transfers are more likely to be spent on durables, suggesting that households face savings and credit constraints. Together, these results suggest that unconditional cash transfers have significant impacts on economic outcomes and psychological well-being.

Keywords

I. Introduction

Unconditional cash transfers (UCTs) have recently received renewed attention as a tool for poverty alleviation in developing countries (Baird, De Hoop, and Özler 2013; Blattman, Fiala, and Martinez 2014). Compared to in-kind transfers, UCTs are attractive because cash is fungible and thus cannot be extramarginal and distortionary; households with heterogeneous needs may be better able to turn cash into long-run welfare improvements than transfers of livestock or skills. UCTs may also have psychological benefits by allowing recipients to choose how to spend money. In addition, UCTs typically have lower delivery

costs than in-kind transfers and are cheaper than conditional cash transfers because no conditions need to be monitored.

On the other hand, UCTs have potential disadvantages from a policy perspective: they might be spent on temptation goods and thereby decrease welfare in the long run; they could lower labor supply due to their income effect (Cesarini et al. 2015); or they could lead to conflict within the family or community (Bobonis, Gonzalez-Brenes, and Castro 2013; Hidrobo, Peterman, and Heise 2016). Relative to conditional cash transfers, they may be inferior in improving the outcomes associated with conditions, but superior in improving other outcomes (Baird, McIntosh, and Özler 2011).

In this article, we shed further light on the impacts of UCTs on important economic and psychological outcomes by studying the program of the NGO GiveDirectly (GD) in Kenya. Between 2011 and 2013, GD sent UCTs of at least US$404 PPP, or at least twice the average monthly household consumption in the area, to randomly chosen poor households in western Kenya using M-Pesa, a cell-phone–based mobile money service.[1] The average transfer amount was $709 PPP, which corresponds to almost two years of per capita expenditure. The GD program is a good laboratory to study the effects of unconditional transfers because existing programs often make relatively small transfers, make large transfers but over a longer period, or target transfers at small business owners. In contrast, GD makes relatively large transfers over a short period of time, targeted at recipients who were chosen simply for meeting a basic means test criterion. In addition, because GD was only beginning to operate in Kenya when the study started, recipients are unlikely to have expected the transfers. We can therefore assess the response of a broad sample of households to large, unanticipated wealth changes.

We study the response of households to these wealth changes using a randomized controlled trial. We carried out a two-stage randomization, one at the village level, resulting in treatment and control villages, and the other at the household level, resulting in "treatment" and "spillover" households in treatment villages, and "pure control" households in control villages. Furthermore, within the treatment group, we randomized the transfer recipient within the household (wife versus husband), the transfer timing (monthly installments over nine months versus one-time lump-sum transfer), and transfer magnitude ($404 PPP versus $1,525 PPP).

This setup allows us to assess the impact of UCTs and address a number of additional questions in the economics literature. First, we document the economic effects of UCTs on consumption (including temptation goods), asset holdings, and income, as well as broader welfare effects on health, food security, education, and female empowerment. We study in detail the effects on psychological well-being, including the stress hormone cortisol. Second, the recipient gender randomization allows us to test whether households are unitary. Third, we use the random assignment of transfer magnitude to ask whether returns to transfers are increasing or decreasing in transfer amount. Finally, the randomization of transfer timing

---

[1] All US$ values are calculated at purchasing power parity, using the World Bank PPP conversion factor for private consumption for KES/US$ in 2012, 62.44. The price level ratio of PPP conversion factor (GDP) to KES market exchange rate for 2012 was 0.5. These figures were retroactively changed by the World Bank after 2013; we use those that were current at the time the study was conducted.

provides evidence on the existence of savings and credit constraints. Because of the large number of outcomes, we address issues of multiple inference by prespecifying the analyses and by using index variables and family-wise error rate correction (FWER).

Nine months after the start of the program, we observe an increase in monthly nondurable expenditure of $36 PPP relative to the spillover group mean of $158 PPP. The treatment effects on alcohol and tobacco expenditure are negative and insignificant, although a lack of power does not allow us to rule out reasonably sized increases. We find a significant increase of $302 PPP in asset holdings, relative to a control group mean of $495 PPP. These investments translate into an increase in monthly revenue from agriculture, animal husbandry, and enterprises of $16 PPP relative to a control group mean of $49 PPP. However, this revenue increase is largely offset by an increase in flow expenses ($13 PPP relative to a control group mean of $24 PPP), and is lower than the returns to capital documented in previous studies (De Mel, McKenzie, and Woodruff 2008; Blattman, Fiala, and Martinez 2014; Fafchamps and Quinn 2015; McKenzie 2015). We find no large effects on health and educational outcomes.

Transfers have a sizable effect on psychological well-being; in particular, we document a 0.16 std. dev. increase in happiness, a 0.17 std. dev. increase in life satisfaction, a 0.26 std. dev. reduction in stress, and a significant reduction in depression (all measured by psychological questionnaires). These results are broadly consistent with those of previous studies on cash transfers (Ozer et al. 2011; Baird, De Hoop, and Özler 2013) and other welfare programs (Bandiera et al. 2013; Ghosal et al. 2013; Banerjee et al. 2015a; see Lund et al. 2011 for a review). Cortisol levels do not show an average treatment effect, suggesting that self-report measures of psychological well-being may be more sensitive to the intervention, or more affected by demand effects (response bias). However, cortisol levels vary across the treatment arms: they are significantly lower when transfers are made to the wife rather than the husband, when they are lump-sum rather than monthly, and when they are large rather than small.

The different treatment arms in the study allow us to address several other questions in the economics literature. First, our design allows us to identify differences in expenditure patterns and other outcomes when transfers are made to the husband versus the wife. We observe few differences between female and male recipient households in consumption, production, and investment decisions, in line with the results of another recent cash transfer experiment (Benhassine et al. 2013). However, because of relatively low power in this cross-randomization, we can pick up only relatively large effects, and thus these results do not argue strongly against findings in other studies that households do not behave in a unitary fashion (Thomas 1990; Duflo and Udry 2004). Second, by randomizing the timing of transfers (monthly versus lump sum), we can ask whether households are both savings and credit constrained. If this were the case, we would expect fewer purchases of expensive assets such as metal roofs among monthly transfer recipients, because the savings constraint would prevent this group from saving their transfer to buy the asset, and the credit constraint would prevent it from borrowing against the promise of the future transfer. We find that indeed this is the case. Finally, we find that the treatment effects for large versus small

transfers are somewhat less than proportional in most categories, suggesting decreasing returns to large transfers overall.

Some of our findings are null results, for which it can be difficult to distinguish between lack of effect and lack of power. This difficulty suggests an innovation in reporting results. For each null finding, we compute the minimum detectable effect size (MDE), that is, the effect that would have been detectable with 80% power at the 5% significance level ex post (Duflo, Glennerster, and Kremer 2007; Appendix Table A.1).[2] This approach provides an intuitive metric to distinguish tightly identified null results from those which fail to reach significance but for which we cannot rule out treatment effects with confidence.

We present two further innovations. First, as is common practice now, our anlayses were specified in a preanalysis plan (PAP).[3] However, while the analyses we report adhere closely to the PAP in general, they deviate occasionally. As a novel approach to increasing transparency, we report all of these deviations, and the reasons for them, in Appendix Table A.3.

Second, we also follow common practice by making public the data and code that produce the results we report in this article. However, it has recently been shown that data and code used in economics papers frequently contains errors, making it difficult for readers to confirm the findings (Chang and Li 2015). We therefore hired two graduate students to audit the data and code for this study. They were compensated on an hourly basis and paid a bonus for any errors they identified. We report the errors they identified and changes they suggested in Online Appendix Section 20. The errors were minor and did not materially change the results and interpretation. We also report which suggested changes we rejected and why.

Our preferred specification relies on within-village treatment effects. For these estimates to be valid, within-village spillovers need to be small. However, one concern regarding identification of spillovers is that there was endogenous selection of pure control households into the survey. The reason for this endogeneity is that these households were chosen based on a thatched-roof criterion, but this criterion was applied at endline rather than at baseline. This fact complicates identification of spillovers, which in turn raises the possibility that our within-village treatment effect estimates are biased. To bound this bias, we resurveyed all metal-roof households in pure control villages to ask when they upgraded to a metal roof. This approach allows us to compute the precise spillover effect of treatment on metal roof ownership. It turns out that this spillover effect is five households, or 1.1% of the sample. We use a number of bounding approaches and find that the resulting bias is small, making our within-village treatment estimates interpretable.

The remainder of the article is organized as follows. Section II describes the GiveDirectly program. Section III summarizes the evaluation design. Section IV presents the impacts of

---

[2]·We are grateful to an anonymous reviewer for this suggestion.

[3]·We wrote two PAPs: one before the first round of analysis, and one before performing additional analyses requested by journal reviewers. Both are available at https://www.socialscienceregistry.org/trials/19/.

the program on all outcomes, including psychological well-being and cortisol levels. Section V concludes.

II. The GiveDirectly UCT Program

GiveDirectly is an international NGO founded in 2009 whose mission is to make unconditional cash transfers to poor households in developing countries. GD began operations in Kenya in 2011 (Goldstein 2013). At the time of the study, eligibility was determined by living in a house with a thatched (rather than metal) roof.[4] Such households were identified through a census conducted with the help of the village elder. After identification, recipient households were visited by a representative of GD, who asked to speak to the transfer recipient in private, collected some demographics, and informed the recipient that they would receive a transfer of KES 25,200 ($404 PPP). Recipients were told that the transfer was unconditional, and they were informed about the transfer schedule (lump sum versus monthly) and the one-time nature of the transfer. Recipients were provided with a Safaricom SIM card and asked to register for the mobile money service M-Pesa.[5] Registration had to occur in the name of the designated transfer recipient. For lump-sum recipients, an initial transfer of KES 1,200 ($19 PPP) was sent on the first of the month following the initial GD visit as an incentive to register.[6]

Withdrawals and deposits can be made at any M-Pesa agent, of which Safaricom operated about 11,000 throughout Kenya at the time of the study. GD estimates the average travel time and cost from recipient households to the nearest M-Pesa agent at 42 minutes and $0.64 (nominal). Withdrawals incur costs between 27% for $2 (nominal) withdrawals and 0.06% for $800 (nominal) withdrawals. GD reports that recipients typically withdraw the entire balance of the transfer upon receipt.

The costs of the GD program at the time of the study amounted to $81 (nominal) per household. Given the proportions of large and small transfers sent at the time, this figure translates to a cost of $113 (nominal) per household for a large transfer ($1,000 nominal), that is, 11%, and $69 (nominal) for a small transfer ($300 nominal), that is, 23%. Of these costs, $50 (nominal) were fixed costs for identification and enrollment of households. Variable costs for foreign exchange and other fees were 6.3% of the transfer amount. Note that in GD's current operating model, only large transfers of $1,000 (nominal) are made, at a cost of $99 (nominal) per transfer.

---

[4] GD had independently established this eligibility criterion as predictive of poverty. We confirm that metal-roof ownership correlates positively with expenditure and asset holdings by comparing our baseline households to a sample of two metal-roof households in every treatment village surveyed at baseline (Online Appendix Table 4).

[5] The treatment is thus a combination of cash transfers and encouragement to register for M-Pesa. We discuss the possible effects of M-Pesa access on outcomes in Section III.D.

[6] GD's operating model has changed since the time of the study. Eligibility is now based not only on a census conducted with a village guide, but is additionally verified by physical back-checks, data back-checks, and crowd-sourced labor to confirm recipient identity and thatched-roof ownership. Transfer amounts are now $1,000 (nominal) per household (corresponding to the "large" transfer amount in the present study), and all eligible households in a village receive transfers.

III. Design and Methods

III.A. Treatment Arms

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 versus lump-sum transfers), and the magnitude of the transfer. The treatment arms were structured as follows.

1. Transfers to the Woman Versus the Man in the Household—Among households with both a primary female and a primary male member, we randomly assigned the woman or the man to be the transfer recipient with equal probability. One hundred ten households had a single household head and were not considered in the randomization of recipient gender.

2. Lump-sum Transfers Versus 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 ($404 PPP). In the lump-sum condition, this amount includes an initial transfer of KES 1,200 ($19 PPP) to incentivize M-Pesa registration, followed by a lump-sum payment of KES 24,000 ($384 PPP). In the monthly condition, the total amount consists of a sequence of nine monthly transfers of KES 2,800 ($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 immediately following the announcement visit by GD, 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 Versus 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 ($1,121 PPP), paid in seven monthly installments of KES 10,000 ($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, and therefore the total transfer amount received by these households was KES 95,200 ($1,525 PPP, $1,000

nominal).[7] The remaining 366 treatment households constitute the "small" transfer group, and received transfers totaling KES 25,200 ($404 PPP, $300 nominal) per household.

These three treatment arms were fully cross-randomized, except that, as noted, 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.

III.B.   Sample Selection and Timing

This study is a two-level cluster-randomized controlled trial. An overview of the design and timeline is shown in Figure I. The selection and surveying of recipient households proceeded as follows (additional details are given in Online Appendix Section 5).

    i.   GD first identified Rarieda, Kenya, as a study district, based on data from the national census. The research team 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 (Online Appendix Table 3). An average of 19% of households per village were surveyed, and an average of 9% received transfers. The transfers sent to villages amounted to an average of 10% of aggregate baseline village wealth (excluding land). A map of treatment and control villages is shown in Online Appendix Figure 1.

    ii.   The research team 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 April–June 2012 in control villages. The census was conducted in the same fashion in treatment and control villages; we address the timing difference in detail in Section IV.B. 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.

    iii.   Following the census, all eligible households completed the baseline survey 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.

    iv.   GD then repeated the census in treatment villages 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

---

[7]·Note 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.

GD's census exercise. We excluded 89 households who completed baseline but were not identified as eligible in the GD census. In Online Appendix Tables 1–2, we show that these households do not differ significantly from the rest of the sample on index variables and demographics. 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.

v. 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 (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). We discuss the possible effects of M-Pesa access in Section II.D.

vi. 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 that did not own one (by reducing the future transfer by the cost of the phone).

For the sample as a whole, transfers were sent between June 2011 and January 2013. Households received the first transfer an average of 4.8 months after baseline and an average of 9.3 months before endline (Online Appendix Figs. 2–3 and Tables 10–13). The last transfer arrived an average of 9.8 months after baseline and 4.4 months before endline. The mean transfer arrived an average of 7.1 months after baseline and 6.9 months before endline.[8] Online Appendix Figure 2 shows histograms of the numbers of surveys and transfers completed in each month; Online Appendix Figure 3 shows histograms for the time elapsed between survey rounds and transfers, including mean/median/minimum/maximum delays between baseline/endline and the first/last/mean/median

---

[8.]The mean transfer date is defined as the date at which half of the total transfer amount to a given household has been sent.

transfers. Online Appendix Tables 10–13 provide summary statistics for the same information, including individual treatment arms.

vii. 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 as "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 bound this bias in Section IV.B. Within treatment villages, treatment and spillover households completed endline on the same day on average, with a nonsignificant difference of 1 day (0.03 month; Online Appendix Table 14). We find a small timing difference in endline date between treatment and control villages, with the latter being surveyed an average of two weeks after treatment villages (0.54 month). In addition, there was some variation across treatment arms in the average delay between the first/mean/last transfer and endline (Online Appendix Tables 10–13). However, in Online Appendix Table 15, we show that endline timing did not correlate with household characteristics. In addition, when we introduce a control variable for survey timing in the main specification, results do not change (Online Appendix Table 16).

III.C. Data Collection

1. Surveys, Biomarkers, and Anthropometrics—In each surveyed household, we collected two distinct modules: a household module, which collected information about assets, consumption, income, food security, health, and education; and an individual module, which collected information about psychological well-being, intrahousehold bargaining and domestic violence, and economic preferences. The two surveys were administered on different (usually consecutive) days. The household survey was administered to any household member who could give information about the outcomes in question for the entire household; this was usually one of the primary members. The individual survey was administered to both primary members of the household, that is, husband and wife, for double-headed households; and to the single household head otherwise. During individual surveys, particular care was taken to ensure privacy; respondents were interviewed by themselves, without the interference of other household members, especially the spouse.

In addition, we measured the height, weight, and upper-arm circumference of the children under five years of age who lived in the household. From a randomly selected subset of (on average) three respondents in each village, we obtained village-level information about prices, wages, and crime to assess possible equilibrium effects of the intervention. A list of

variables collected is presented in Online Appendix Section 1, and all questionnaires are available at https://www.socialscienceregistry.org/trials/19.

In addition to questionnaire measures of psychological well-being, we obtained saliva samples from all respondents, which were assayed for the stress hormone cortisol. Measuring cortisol has several advantages over other outcome variables. First, it is an objective measure and not prone to survey effects such as social desirability bias (Zwane et al. 2011; Baird and Özler 2012), and it is easy to measure in field settings.[9] Second, cortisol is a useful indicator of both acute stress (Kirschbaum, Strasburger, and Langkrär 1993; Ferracuti et al. 1994) and more permanent stress-related conditions such as major depressive disorder (Holsboer 2000; Hammen 2005). Third, cortisol is a good predictor of long-term health through its effects on the immune system.[10]

We obtained two saliva samples from each respondent, one at the beginning of the individual survey and the other at the end, using the Salivette sampling device (Sarstedt, Germany). The Salivette has been used extensively in psychological and medical research (Kirschbaum and Hellhammer 1989), and more recently in developing countries in our own work and that of others (Fernald and Gunnar 2009; Chemin, Haushofer, and Jang 2016). It consists of a plastic tube containing a cotton swab, on which the respondent chews lightly for two minutes to fill it with saliva. Due to the noninvasive nature of this technique, we encountered no apprehension among respondents. The saliva samples were labeled with barcodes and stored in a freezer at −20 °C, and were later centrifuged and assayed for salivary free cortisol using a standard radioimmunoassay on the cobas e411 platform at Lancet Labs, Nairobi.

Cortisol levels were analyzed as follows: We first obtained the average cortisol level in each participant by averaging the values of the two samples. Because cortisol levels in population samples are usually heavily skewed, it is established practice to log-transform them before analysis; we follow this standard approach here. Salivary cortisol is subject to a number of confounds; in particular, it is affected by food and drink, alcohol and nicotine, medications, and strenuous physical exercise. Cortisol levels also follow a diurnal pattern—they rise sharply in the morning, and then exhibit a gradual decline throughout the rest of the day. To control for these confounds but at the same time avoid the risk of cherry-picking control variables, we consider two measures of cortisol in the analysis. First, we use the log-

---

[9.]Cortisol can be measured noninvasively in saliva, where it is a good indicator of levels in the blood (Kirschbaum and Hellhammer 1989); it is stable for several weeks, even without refrigeration; and commercial radioimmunoassays for analysis are widely available at relatively low cost. Strictly speaking, it is possible to "lie" about cortisol levels, in the sense that they can be intentionally manipulated through food, caffeine, or alcohol intake, as well as physical exercise. However, three factors make it unlikely that our respondents undertook such manipulation. First, for it to systematically affect our results, our participants would have to have intimate knowledge of the environmental and physiological factors that affect cortisol levels, which we deem unlikely. Second, a group of participants would have to concertedly use this knowledge, in a coordinated fashion, to attempt to bias our results. Third, this manipulation would have to be outside the scope of our control variables, which include all the factors that commonly affect cortisol levels, or participants would have to systematically lie about certain control variables. Given the fact that our respondents largely appeared unaware of what cortisol was, much less how physiological and environmental factors affected it, we judge it as highly unlikely that participants systematically and intentionally manipulated cortisol levels.

[10.]Cortisol exerts a direct and broadly suppressive effect on the immune system; in particular, it suppresses pro-inflammatory cytokines such as interleukin-6 and interleukin-1 (Straub 2006;Wilckens 1995).Chronic elevations of cortisol have the opposite effect, leading to permanent mild elevations of cytokine levels (Kiecolt-Glaser et al. 2003). These cytokine elevations then contribute directly to disease onset and progression, for example in atherosclerosis and cancer (Ross 1999; Steptoe et al. 2001, 2002; Coussens and Werb 2002; Aggarwal et al. 2006). Thus, permanently high cortisol is physiologically damaging, over and above its psychological significance.

transformed raw cortisol levels without the inclusion of control variables. Second, we construct a "clean" version of the raw cortisol levels, which consists of the residuals of an OLS regression of the log-transformed cortisol levels on dummies for having ingested food, drinks, alcohol, nicotine, or medications in the two hours preceding the interview; for having performed vigorous physical activity on the day of the interview; and for the time elapsed since waking (rounded to the next full hour). We include both the raw and clean versions of the cortisol variable in the analysis. The resulting estimates for the two versions are nearly identical.

2.   Main Outcome Variables—We compute the following index variables (further detail is given in Appendix Table A.2 and Online Appendix Section 2).

i.   "Value of nonland assets" is the total value (in 2012 US$PPP) of all nonland assets owned by the household, including savings, livestock, durable goods, and metal roofs.

ii.   "Nondurable expenditure" is the total monthly spending (in 2012 US$ PPP) on nondurables, including food, temptation goods, medical care, education expenditures, and social expenditures.

iii.   "Total revenue" is the total monthly revenue (in 2012 $PPP) from all household enterprises, including revenue from agriculture, stock and flow revenue from animals owned by the household, and revenue from all nonfarm enterprises owned by any household member.

iv.   The food security index is a standardized weighted average of the (negatively coded) number of times household adults and children skipped meals, went whole days without food, had to eat cheaper or less preferred food, had to rely on others for food, had to purchase food on credit, had to hunt for or gather food, had to beg for food, or went to sleep hungry in the preceding week; a (negatively coded) indicator for whether the respondent went to sleep hungry in the preceding week; the (positively coded) number of times household members ate meat or fish in the preceding week; (positively coded) indicators for whether household members ate at least two meals per day, ate until content, had enough food for the next day, and whether the respondent ate protein in the last 24 hours; and the (positively coded) proportion of household members who ate protein in the last 24 hours, and proportion of children who ate protein in the last 24 hours.[11]

v.   The health index is a standardized weighted average of the (negatively coded) proportion of household adults who were sick or injured in the last month, the (negatively coded) proportion of household children who were sick or injured in the last month, the (positively coded) proportion of sick or injured family members for whom the household could afford treatment, the (positively coded) proportion of illnesses for which a doctor was consulted, the (positively coded) proportion of newborns who were vaccinated, the (positively coded) proportion

---

[11].All weighted standardized averages were computed using the approach described in Anderson (2008).

of children below age 14 who received a health checkup in the preceding six months, the (negatively coded) proportion of children under age 5 who died in the preceding year, and a children's anthropometrics index consisting of body mass index, height-for-age, weight-for-age, and upper-arm circumference relative to WHO development benchmarks.

vi. The education index is a standardized weighted average of the proportion of household children enrolled in school and the amount spent by the household on educational expenses per child.

vii. The psychological well-being index is calculated separately for the primary male and primary female in the household and is a standardized weighted average of their (negatively coded) scores on the CESD scale (Radloff 1977), a custom worries questionnaire (negatively coded), Cohen's stress scale (Cohen, Kamarck, and Mermelstein 1983; negatively coded), their response to the the World Values Survey happiness and life satisfaction questions, and their log cortisol levels adjusted for confounders (negatively coded).

viii. The female empowerment index is reported for the household's primary female only, and consists of a standardized weighted average of a measure of two other indexes, a violence and an attitude index. The violence index is a weighted standardized average of the frequency with which the respondent reports having been physically, sexually, or emotionally abused by her husband in the preceding six months; the attitude index is a weighted standardized average of a measure of the respondent's view of the justifiability of violence against women, and a scale of male-focused attitudes.

III.D. Integrity of Experiment

1. Baseline Balance—We test for baseline differences between treatment and control groups using the following specification.

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \epsilon_{vhiB}. \tag{1}$$

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 spillover households.[12] Village-level fixed effects are captured by α_v. T_{vh} is a treatment indicator that takes the 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, that is, the household. In addition to this standard inference, we compute FWER-corrected *p*-

---

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

values across the set of index variables (Anderson 2008), but not across specifications for the overall treatment effect and the different treatment arms. 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.

The results of this estimation for our index variables are shown in Table I. In column (2) we report the difference in baseline outcomes between control and treatment households in treatment villages. The results are largely insignificant, suggesting that the treatment and control groups did not differ at baseline. The only significant difference between treatment and control households appears in income from self-employment, where treatment households have a $33 PPP lower income relative to the control mean of $85 PPP (39%) at baseline. This difference is significant at the 10% level, but does not survive FWER correction for multiple inference. Note that it goes in the conservative direction.

In columns (3)–(5) of Table I, we report the difference in baseline outcomes between treatment arms, restricting the sample to treatment villages. Column (3) reports differences between treatment households in which transfers were made to the female versus the male in the household, analyzed as follows:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\mathrm{F}} + \beta_2 T_{vh}^{\mathrm{W}} + \beta_3 S_{vh} + \epsilon_{vhiB}. \tag{2}$$

Here, the variables T_{vh}^{X} are indicator functions that specify whether the transfer recipient is female $\left(T_{vh}^{\mathrm{F}}\right)$ 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) $\left(T_{vh}^{\mathrm{W}}\right)$. 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. Column (3) of Table I reports β_1, that is, the difference in baseline outcomes between female and male recipient households.

In column (4), we report the differential effect of monthly versus lump-sum transfers, analyzed as follows:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\mathrm{MTH}} \times T_{vh}^{\mathrm{S}} + \beta_2 T_{vh}^{\mathrm{L}} + \beta_3 S_{vh} + \epsilon_{vhiB}. \tag{3}$$

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, and column (4) of Table I reports β_1, that is, the difference in baseline outcomes between monthly and lump-sum recipient households.

Finally, in column (5), we report the effect of receiving large compared to small transfers, using the following specification:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{L} + \beta_2 S_{vh} + \epsilon_{vhiB}. \tag{4}$$

Here, T_{vh}^{L} is an indicator variable for having been assigned to receiving large transfers. Thus, column (5) reports β_1, the difference in baseline outcome measures between households receiving large transfers and households receiving small transfers.

Several significant differences appear between treatment arms, notably for food security between monthly and lump-sum transfer recipients, where monthly transfer households have a 0.25 std. dev. higher score on the food security index than lump-sum transfer households at baseline, and the health index, which is 0.26 std. dev. higher in households where the female receives the transfer. Both of these differences are significant after FWER correction, though the difference in food security is only significant at the 10% level. None of the other treatment arm differences are significant after FWER correction.

2.   Compliance—Due primarily to registration issues with M-Pesa, 18 treatment households had not received transfers at the time of the endline, and thus only 485 of the 503 treatment households were in fact treated. We deal with this issue by using an intent-to-treat approach, and consider all households assigned to receive a transfer as the treatment group, regardless of whether they had received a transfer at the time of the endline survey.

3.   Attrition—We had low levels of attrition; overall, 940 of 1,008 baseline households (93.3%) were surveyed at endline. In the treatment group, 471 of 503 baseline households (94%) were surveyed at endline, and in the spillover group, 469 of 505 (93%). These low levels of attrition suggest that the program did not cause a large number of households to leave the village or to reform. Detailed attrition analyses are shown in Online Appendix Section 8. First, a regression of the attrition dummy on the treatment dummy shows no difference in the likelihood of attrition between the treatment and control groups (Online Appendix Table 6). Second, a regression of our main index variables on the attrition dummy reveals no significant overall difference in outcomes at baseline between attrition and nonattrition households (Online Appendix Table 7). Third, a regression among attrition households of our index variables on the treatment dummy shows that there were no differences in outcomes between attrition households that had been assigned to the treatment and the control condition (Online Appendix Table 8). Finally, bounding the treatment effects on the index variables using Lee bounds (Lee 2009) reveals minimal differences between upper and lower bounds of the treatment effects (Online Appendix Table 9). Thus, attrition is unlikely to have biased the results reported below.

4.   Effects of M-Pesa Access—As described already, our treatment entails not only cash transfers but also provision of a SIM card and an incentive to register for M-Pesa. Together with recent evidence on the consumption smoothing and savings effects of access to M-Pesa and similar savings technologies (Dupas and Robinson 2013b; Jack and Suri 2014), this fact raises the possibility that our economic effects may be partly driven by M-Pesa access. However, we find small effects of treatment on M-Pesa use, reported in Online Appendix Section 17: treatment households save an extra $3 PPP in M-Pesa compared with control households and receive an extra $9 PPP per month in remittances, but do not show

an increase in outgoing remittances (Online Appendix Table 33). These effects are small compared to (for example) the $36 PPP monthly increase in consumption among treatment households. In line with these findings, Aker et al. (2016) find few differences in outcomes when cash transfers are delivered manually versus through mobile money in Niger. They do find differential effects on food security and female empowerment, suggesting that in our study, effects on these outcomes might be partially mediated by access to M-Pesa. However, in our view it is unlikely that the delivery method changed expenditure patterns in our setting, since anecdotally households withdrew the money immediately after receiving it.

III.E.   Data Analysis

1.   PAPs—A goal of this study was to provide a comprehensive picture of the impacts of unconditional cash transfers on households. We therefore collected a large number of outcomes and endeavor in this article to report the full breadth of the evidence. To ensure no cherry-picking of results from these many outcomes, we wrote a PAP for this study, which is published and timestamped at https://www.socialscienceregistry.org/trials/19 (Casey, Glennerster, and Miguel 2012; see also Rosenthal 1979; Simes 1986; Horton and Smith 1999). In the PAP, we specify the variables to be analyzed, the construction of indexes, our approach to dealing with multiple inference, the econometric specifications to be used, and the handling of attrition. The reduced-form analyses and results reported in this article correspond to those outlined in the PAP, with the exception of the restriction of the sample to thatched-roof households at endline when identifying spillover effects, to account for the time delay in applying the thatched-roof criterion to the pure control group. In addition, after article submission, we wrote a second PAP to address reviewer comments. This second PAP is also available at https://www.socialscienceregistry.org/trials/19. On a few occasions, the analyses we report in this article deviate in a minor fashion from those specified in the PAPs. We report these deviations, and the reasons for them, in Appendix Table A.3.

2.   Reduced-Form Specifications—Our basic treatment effects specification to capture the impact of cash transfers is:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \epsilon_{vhiE}. \tag{5}$$

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 IV.B. 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}).[13] All other features are as in equation (1).

[13.]Jones (1996) shows that under some circumstances this approach can yield biased estimates; however, we had no missing baseline observations in the majority of outcome categories, and only 54, 51, and 73 in the education, psychological well-being, and female empowerment indexes, making it unlikely that such bias strongly affected our estimates.

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 versus the male in the household is captured by the following model, restricting the sample to treatment villages and denoting spillover households with S_{vh}:

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\text{F}} + \beta_2 T_{vh}^{\text{W}} + \beta_3 S_{vh} \\ + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \epsilon_{vhiE}. \tag{6}$$

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

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\text{MTH}} \times T_{vh}^{\text{S}} + \beta_2 T_{vh}^{\text{L}} + \beta_3 S_{vh} \\ + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \epsilon_{vhiE}. \tag{7}$$

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

$$y_{vhiE} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\text{L}} + \beta_2 S_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \epsilon_{vhiE}. \tag{8}$$

All restrictions and other features are as described in Section III.D.

3.   Accounting for Multiple Comparisons—Due to the large number of outcome variables in the present study, false positives are a potential concern when conventional approaches to statistical inference are used. We employ two strategies to avoid this problem, following broadly the approaches of Kling, Liebman, and Katz (2007), Anderson (2008), and Casey, Glennerster, and Miguel (2012).

First, we compute standardized indexes for several main groups of outcomes and choose focal variables of interest for others (all specified in the PAP), as described already. Second, even after collapsing variables into indexes and choosing focal variables of interest for each group of outcomes, we are still left with multiple indexes, creating the need to further control the probability of Type I errors. To this end, we control the FWER (Westfall, Young, and Wright, 1993; Efron and Tibshirani 1993; Anderson 2008; Casey, Glennerster, and Miguel 2012) across the treatment coefficients on the indexes for our main outcome groups, that is assets, consumption, income, psychological well-being, education, food security, health, and female empowerment. As specified in our PAP, we apply this correction to the index variables only; when discussing individual variable results within particular outcome groups, we use conventional significance levels. We use this approach because the purpose of studying individual variables within the outcome groups is to understand mechanisms, rather than single out particular variables for general conclusions. Further detail on controls for multiple inference is given in Online Appendix Section 3.

4.   Minimum Detectable Effect Sizes—For null results, we report the minimum detectable effect size (MDE) with 80% power at a significance level of 0.05. It is given by

$$MDE = (t_{1-\kappa} + t_{\frac{\alpha}{2}}) \times \frac{\sigma}{\sqrt{NP(1-P)}},$$

where $t_{1-\kappa}$ is the value of the t-statistic required to obtain 80% power, $t_{\frac{\alpha}{2}}$ is the critical t-value required to achieve a significance level of 0.05, P is the fraction of the sample that were treated, and $\frac{\sigma}{\sqrt{NP(1-P)}}$ is the standard error of the treatment coefficient. With $P = 0.5$, $t_{1-\kappa} = 0:84$, and $t_{\frac{\alpha}{2}} = 1.96$, this expression simplifies to a simple multiple of the standard error of the treatment coefficient, $SE(\hat{\beta})$:

$$MDE = 2.8 \times SE(\hat{\beta}).$$

We use this formula to compute the *MDE* for null results. In the case of outcome variables with a nonzero control group mean (such as monetary variables), we additionally report the *MDE* as a proportion of the control group mean.

IV. Results

IV.A. Direct Effects on Index Variables

1. Overall Impacts—Table II shows the main results of the program for the index variables selected in the PAP. Column (1) shows the means and standard deviations in the spillover group. The treatment effect for these variables is shown in column (2), estimated using equation (5). Standard errors are shown in parentheses, and the bootstrapped FWER p-values (10,000 iterations) in brackets. The last row of the table reports the joint significance of all coefficients in the corresponding column, using SUR.

We find statistically significant and economically meaningful impacts of cash transfers across the majority of outcomes measured by our indexes, including assets, consumption, food security, revenue from self-employment, and psychological well-being. Overall, the joint significance of the treatment effects across outcomes has a p-value of less than .005. Household consumption of nondurable goods is significantly higher in the treatment group than in the control group ($36 PPP a month, or 23% of the control group consumption).[14] The effect is statistically significant at the 1% confidence level according to both standard and FWER p-values. Concomitantly, we observe significant improvement in the food security index (0.26 std. dev.). Households invest part of the transfers: the value of nonland assets increased by $302 PPP on average; this represents 61% of the control group mean of $495 PPP and 43% of the average transfer.[15] The effect is statistically different from 0 at the 1% confidence level according to both standard and FWER-corrected p-values. For monthly agricultural and business income, the point estimate on the treatment effect shows a $16 PPP increase. This is an increase of 33% over the control group mean of $49 PPP, and on an annual basis, it represents 27% of the average transfer amount. We see no improvement in health, education, or female empowerment when comparing treatment and control

---

[14.]All monetary variables were top-coded at 99% and coded linearly. However, because these outcome variables are skewed even after top-coding, we additionally present log specifications in Online Appendix Tables 42–48. In doing so, we use the inverse hyperbolic sine transform to deal with zeros (Burbidge, Magee, and Robb, 1988; MacKinnon and Magee 1990; Pence 2006), which transforms each outcome variable as follows:

[15.]Twenty-eight percent of the treatment group received a transfer of KES 95,200 ($1,525 PPP), while the remaining 72% received KES 25,200 ($404 PPP). The average transfer was $709 PPP (note that this average is calculated using unrounded figures and therefore differs slightly from 1; $525 \times 0:28 + 404 \times 0:72$).

households within the same villages. Appendix Table A.1 reports MDEs for the main outcome variables, and shows that we were powered to detect relatively small effect sizes for these outcomes (0.17 std. dev. for health, 0.16 for education, and 0.20 for female empowerment). The lack of difference between treatment and spillover households with respect to female empowerment is due to both groups reporting higher female empowerment than pure control households; we discuss this result later. The lack of overall effects on health or education may be due to the relatively short-term nature of the follow-up in this study; because these outcomes move on longer time scales, it is possible that they may show changes in longer-term data.

$$y' = \ln\left(y + \sqrt{y^2 + 1}\right). \tag{9}$$

The results we find in the log specifications are similar to those reported for the linear specifications reported here.

Online Appendix Table 35 shows only minor differences in the estimates of the treatment effects when baseline controls are included; none of the significant results become nonsignificant or vice versa. Thus, baseline covariates do not affect our results strongly. In addition, the magnitudes and significance levels obtained in this within-village analysis are broadly similar to those found when comparing treatment to pure control households (Online Appendix Table 38), with three exceptions. First, the treatment effect on assets is larger when estimated across villages than within villages. Second, due to a large within-village spillover effect of 0.21 std. dev., the coefficient on the female empowerment index is small and not significant in the within-village comparison (MDE 0.20 std. dev.), but is large (0.20 std. dev.) and significant at the 5% level in the across-village comparison. Finally, the increase in revenue is not significant when estimated across villages (*MDE* $16.46 PPP, 34% of control group mean).

2. Effects of Treatment Arms—We now discuss the three subtreatments: transfers to the primary female versus the primary male in the household, monthly versus lump-sum transfers, and large versus small transfers.

Column (3) in Table II reports the coefficients and standard errors comparing female to male recipient households on the index variables. With the exceptions of psychological well-being and female empowerment, significant at the 10% level and further discussed below, none of the differences between the treatment effects for transfers to a woman and those to a man are statistically significant at conventional significance levels. Thus, we find little evidence that providing cash transfers to women versus men differentially affects outcomes, broadly in line with the findings of Benhassine et al. (2013).[16] However, we note that this lack of significant differences may result from lower power in the comparison between male and female recipient households, because single-headed households are excluded in this analysis, reducing the sample size. Our MDEs ranged from 0.21 to 0.29 std. dev. for standardized

---

[16] Another existing study randomizes the recipient gender of UCTs, but the results are not available yet (Akresh, de Walque, and Kazianga 2013); Baird, McIntosh, and Özler (2011) independently randomize transfer amounts to girls and their parents, finding few differences in schooling and pregnancy outcomes.

outcomes, and between 18% and 61% of the control group mean for monetary outcomes (Appendix Table A.1). We do observe a trend in the point estimates, suggesting that transferring cash to the primary male in the household leads to a larger impact on standard measures of economic welfare, namely, assets and consumption, while transferring cash to the primary woman in the household improves outcomes most likely to benefit children, that is food security, health, and education, as well as psychological well-being and female empowerment. It is possible that a more highly powered study would observe significant differences in these outcomes.

Results comparing monthly to lump-sum transfers are shown in column (4) of Table II. The joint significance across outcomes is $p < .05$, suggesting that monthly and lump-sum transfers have significantly different effects on our outcomes. In the individual variables, we find that monthly payments increase food security by 0.26 std. dev. relative to lump-sum payments, and that lump-sum transfers lead to higher levels of asset holdings than monthly transfers; both effects are statistically significant at conventional levels but do not survive FWER correction. We discuss the finding that larger expenditures on assets are possible with lump-sum than with monthly transfers in more detail in Section IV.E.

Finally, column (5) of Table II compares large to small transfers. We find large and highly significant differences between large and small transfers, all in the direction of "better" outcomes for large transfers. The joint significance across outcomes has a p-value of less than .005. Regarding individual outcomes, most prominently, the increase in asset holdings resulting from the large transfer is approximately twice as large as that for the small transfer. The differences between the subgroups on these outcomes are statistically significant in terms of both conventional and FWER-adjusted p-values. In addition, we find that larger transfers improve the psychological well-being of household members to a greater extent than small transfers; this difference is also significant in terms of both standard and FWER-adjusted p-values. Finally, we observe an additional increase in female empowerment for large transfers, significant at the 5% level using conventional p-values, but not FWER-corrected inference. These results are broadly consistent with those reported by Baird, McIntosh, and Özler (2011), who find increases in school enrollment rates and decreases in marriage incidence with increasing unconditional cash transfers to parents in Malawi.

Note that in both columns (5) and (6), the coefficients are jointly significant using SUR despite the fact that few of them survive FWER correction. This apparent discrepancy results from the fact that FWER correction is more conservative than joint testing using SUR.

IV.B. Spillovers to Other Households and Equilibrium Impacts

For the results reported in Table II to provide an unbiased estimate of the treatment effect, within-village spillovers of treatment on nonrecipient 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:

$$y_{vhiE} = \beta_0 + \beta_1 S_{vh} + \epsilon_{vhiE}. \tag{10}$$

The sample includes only nontreatment households; 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 (Pepper 2002; Cameron, Gelbach, and Miller 2011). 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}.

The results of this analysis are reported in column (1) of Table III. The spillover effects are generally small and not significant, with one exception: We observe an increase of 0.21 std. dev. in the female empowerment index among the control group in treatment villages. This increase is significant at the 5% level using conventional p-values. Together with a non-significant direct treatment effect of −0.01 std. dev. on this measure, this spillover effect suggests that the treatment group shows a significant increase in female empowerment relative to the pure control group, which we confirm in Online Appendix Table 38. However, since we do not have a good theory for why spillover effects might occur in female empowerment, we do not offer an interpretation of this result at this stage.

More generally, however, we note that most of our spillover effect estimates are relatively precisely measured null effects. This finding alleviates the concern that we have low statistical power to detect spillover effects. The average standard error for the standardized variables is 0.08, which implies that the detectable effect size at a 5% significance level and 80% power was 0.22 std. dev. Thus, we can rule out small spillover effects with relatively high confidence.

1. Are Spillover and Pure Control Households Comparable?

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. In the following, we bound the potential bias arising from this problem; a formal treatment of the problem, and the identifying assumptions for the different approaches, are reported in Online Appendix Section 11.

As a first test of whether spillover and pure control households are comparable, we ask whether they differ significantly on immutable characteristics. Across a number of such variables (respondent age, marital status, number of children, household size, and education level), we find none that differ between the two groups (Online Appendix Table 25). Second, we ask whether roof upgrade can be predicted based on baseline covariates in the spillover group. We find relatively weak predictive power across a large number of variables (Online Appendix Table 26). Together, these findings suggest that households that upgrade are similar to ones that do not, and that the full spillover sample (including households that

upgrade and those that do not) is comparable to the pure control sample. To provide further evidence for this claim, we estimate the spillover effect with control variables in column (2) of Table III. The results are broadly similar to those obtained without controls, with somewhat larger negative point estimates for the asset and expenditure spillovers, and the point estimate on expenditure significant at the 10% level. Again, the most salient spillover effect is that on domestic violence, which remains at 0.21 std. dev.

The negative spillover effect on expenditure contrasts with evidence from conditional cash transfer programs (Angelucci and De Giorgi 2009), and with theory suggesting that households should insure each other (Townsend 1994). However, we note that this effect is small in magnitude and marginally significant. It nevertheless raises the possibility that the within-village treatment effect on expenditure may be overestimated. To test this question directly, Online Appendix Table 38 compares the treatment effect on expenditure when it is estimated within villages (i.e., treatment versus spillover households) and across villages (i.e., treatment versus pure control households). In both cases the effect is large and highly significant.

2.   Restricting the Sample to Households with Thatched Roofs at Endline, and Precise Estimation of the Spillover Effect on Metal-Roof Ownership—Our main approach to improve identification of the spillover effect is to estimate it for only those households that still have thatched roofs at endline. To see why this approach is attractive, consider the *potential* metal-roof upgrade decisions of the households in the spillover and pure control groups. Under a standard monotonicity assumption, those spillover households that have metal roofs at endline are either always takers (i.e., they would have upgraded to a metal roof regardless of spillover versus pure control status) or compliers (i.e., they upgraded to metal roofs because they were in the spillover group). Spillover households that still have thatched roofs at endline are never takers.[17] In the pure control group, (nonsurveyed) households with metal roofs at endline are always takers, and households with thatched roofs are either compliers or never takers. Thus, spillover and pure control households with thatched roofs at endline are comparable if the proportion of compliers is small. In the extreme, if it is 0, the spillover effect for never takers is perfectly identified by the comparison of these two groups.

Importantly, we can find out how many such households there are by obtaining a precise estimate of the magnitude of the spillover effect of the cash transfers on metal-roof ownership. In September 2015, we returned to all households with metal roofs in pure control villages ($N = 3,356$) to ascertain when they upgraded to a metal roof. Households that upgraded between April 2011 and June 2012 should originally have been eligible for participation in the study, but were excluded because of the late application of the thatched-

---

[17.]The monotonocity assumption requires that there be no defiers in the sample. Is this assumption justified? In our view, the only plausible reason for control households to refrain from upgrading their thatched roofs to metal is to remain eligible for possible future transfers from GD. However, control households in treatment villages were credibly told by GD that they would not receive cash transfers. The no-defier assumption is therefore reasonable in our setting. In addition, as detailed in Section 11.4 of the Online Appendix, the identification of the metal-roof spillover effect is also valid under any of three alternative assumptions, namely (i) that the potential outcomes for compliers, never takers, and defiers are the same; (ii) that the proportion and potential outcomes for compliers and defiers are the same; and (iii) that the potential outcomes for compliers and a portion of the defiers are the same (Angrist, Imbens, and Rubin 1996; de Chaisemartin, forthcoming).

roof criterion. We identified 170 such households. Using the same algorithm originally used to select pure control households, we calculate that 78 of these households should have originally been included in the sample. We can now compare the upgrade rates in treatment and pure control villages. Since there were 432 pure control households in the original study, the upgrade rate from baseline to endline in pure control villages is $\frac{78}{432+78} = 0.153$.

Similarly, since there were a total of 469 spillover households at endline, of which 77 had metal roofs the upgrade rate among spillover households was $\frac{77}{469} = 0.164$. Applying the upgrade rate of 0.153 in pure control villages to these spillover households, we would predict $0.153 \times 469 = 72$ metal roofs in the spillover group at endline. In actuality, we observe 77 metal-roof households. The treatment therefore had a spillover effect on metal-roof ownership of $77 - 72 = 5$ households.

We take two approaches to the bias arising from these five households. The first is to ignore it: with 5 households out of 469, that is, 1.1%, the spillover effect of transfers on metal roof ownership is negligible, and therefore restricting the sample to households that still have thatched roofs at endline identifies the spillover effect rather well. The results of this analysis are shown in columns (3) (without controls) and (4) (with controls) of Table III. Again, we find small and largely nonsignificant spillover effects, except for female empowerment and a marginally significant effect on expenditure. And again, comparison of the within-village and across-village treatment effects on these variables (Online Appendix Table 38) shows that the spillover effects do not materially change the magnitude and significance of the direct treatment effects. Columns (5) and (6) of Table III report p-values for the comparison of the spillover estimates when the sample is versus is not restricted to thatched-roof households. We find significant differences for only one variable, assets, suggesting that broadly, the restriction to households with thatched roofs at endline did not affect the results much.

The second approach is to bound the spillover effect using worst-case assumptions about the potential outcomes of the metal-roof spillover households, as implemented by Lee (2009) and Horowitz and Manski (1995).[18] Results are shown in columns (7)–(10) of Table III. Both the upper bounds and lower bounds are small and nonsignificant, the only exceptions being a significant lower Lee bound on the education index (5% level), and the female empowerment results discussed earlier.

3. Differential Attrition among Spillover and Pure Control Households?—A
further potential source of bias in our estimation of the spillover effects is that because of the late selection of the pure control households, this sample might have differential attrition relative to the spillover households. However, it is unlikely that such attrition biased our effect size estimates. First, Table 7 in the Online Appendix shows that households which attrited from the spillover and treatment groups were not systematically different from other households. If the process driving attrition was similar for spillover and pure control

---

[18]·For the latter approach, in principle the theoretical upper and lower bounds of the support of the outcome variable should be used to impute missing observations; for practical purposes, following the suggestion of Lee (2009), we use empirically determined bounds at the 5th and 95th percentiles of the support of the outcome variable.

households, this fact makes it unlikely that the comparison of spillover and pure control households was biased due to attrition. Second, the assumption that the process driving attrition was in fact similar for the two types of households is supported by the fact that attrition in treatment and spillover households was very similar (6% and 7%, respectively; Online Appendix Table 6), and that attriting treatment households were similar to attriting spillover households in terms of baseline characteristics (Online Appendix Table 8). Given this fact, it is even more likely that the attrition process was also similar for spillover and pure control households, because the spillover group is likely more comparable to the pure control group than to the treatment group. Together, these findings make it unlikely that the comparison of spillover and pure control households is significantly affected by attrition.

4. Within-Village Spillovers Based on Treatment Intensity—Another approach to assess the magnitude of within-village spillovers—one that is independent of levels of attrition in the pure control group—is to use random variation across treatment villages in the mean transfer amount to the village. Variation in this variable comes from the fact that as a consequence of randomizing the large and small transfers across villages, some villages received larger average transfers than others. Using this approach, we recently reported negative within-village spillovers of the transfers on life satisfaction (Haushofer, Reisinger, and Shapiro 2015). However, this finding is unlikely to bias the within-village treatment estimates we report here, for two reasons.

First, in Haushofer, Reisinger, and Shapiro (2015), we use an identification strategy based on differences in the average village-level wealth increase across treatment villages, rather than a comparison between treatment and pure control villages: we compare treatment villages in which average wealth increased only slightly to others in which average wealth increased more significantly, and find differences in life satisfaction between them. However, in that paper and the present one, there is no spillover effect when comparing spillover to pure control households. Importantly, the integrity of the within-village treatment estimates that are the focus of the present article relies only on this across-village spillover effect being small. This is the case in this article and in Haushofer, Reisinger, and Shapiro (2015).

Second, in Haushofer, Reisinger, and Shapiro (2015), we only find evidence of externalities on one variable, life satisfaction, among several indicators of psychological well-being. In our view it is unlikely that this effect would have affected the treatment effects on the other outcome variables presented here; although there is evidence for an influence of happiness on productivity (Oswald, Proto, and Sgroi 2009), it is small, has only been identified in the lab, and to our knowledge has not been found for life satisfaction. In line with this argument, when we analyze spillover effects for our set of index variables using the treatment intensity approach, we do not find significant spillover effects, as shown in Online Appendix Table 27. Thus, our estimates of the within-village treatment effects are unlikely to be distorted by spillovers as identified by variation in mean transfer amount across villages.

5. Village-Level Equilibrium Effects—To investigate whether cash transfers had equilibrium effects (on prices, wages, etc.) at the village level, we collected village-level outcomes from multiple individuals in both treatment and control villages. Specifically, a random subset of (on average) three respondents per village were surveyed about prices for a

standard basket of foods and other goods, wages, and crime rates. Related variables were combined into summary indexes as described in Online Appendix Section 2.2. We regress average village-level outcomes at endline ($\bar{y}_{vE}$) on an indicator variable for whether village v is a treatment village:

$$\bar{y}_{vE} = \beta_0 + \beta_1 T_v + \epsilon_{vE}. \tag{11}$$

We present results in Online Appendix Table 149 (with detailed results in Online Appendix Tables 150–159). We find no significant village-level effects, except for a marginally significant effect on the index of nonfood prices. We therefore conclude that the transfers had little effect on village-level outcomes. This is not to rule out this possibility in principle, since only a relatively small proportion of households were treated in each village, and our MDEs ranged from 0.48 to 0.67 std. dev.[19]

In sum, across specifications and bounding approaches, the spillover effects are small in magnitude and rarely significant. They are thus unlikely to materially affect the within-village treatment effects we report in the main tables. In addition, as described above and shown in Online Appendix Table 38, the within-village treatment effects are similar to the across-village treatment effects, in terms of both magnitude and statistical significance.

IV.C. Psychological Well-Being and Cortisol

1. Overall Effects—A central goal of this study was to assess in detail the effects of UCTs on psychological well-being and levels of the stress hormone cortisol. We had hypothesized that cash transfers would lead to an increase in psychological well-being, specifically to a reduction in stress and cortisol levels (Haushofer and Fehr 2014). Overall, the transfers indeed led to a large and significant improvement in psychological well-being; the treatment effect on the psychological well-being index (a standardized weighted average of the clean cortisol levels, worries, stress, depression, happiness, and life satisfaction variables) is 0.26 std. dev., significant at the 1% level according to both standard and FWER-adjusted *p*-values (Table II). Table IV investigates this effect in more detail and shows that it stems mainly from a 0.16 std. dev. increase in happiness scores (measured by the World Value Survey [WVS] question on happiness), a 0.17 std. dev. increase in life satisfaction (also from the WVS), a 0.26 std. dev. reduction in stress (measured by the four-item version of Cohen's Stress Scale, Cohen, Kamarck, and Mermelstein 1983), a 1.2-point reduction in scores on the CESD depression questionnaire (Radloff 1977), a 0.13 std. dev. reduction in self-reported worries (measured using a custom worries scale), and a marginally significant increase in optimism (measured by Scheier's Life Orientation Test [Revised]; Scheier, Carver, and Bridges 1994).[20] That an exogenous reduction in poverty causes significant reductions in stress and depression, and increases in happiness and life satisfaction, lends support to the hypothesis that poverty alleviation has psychological benefits.

---

[19].Another possible explanation for this null finding is that effects at the level of the local economy are larger for in-kind transfers than cash transfers (Currie and Gahvari 2008; Cunha 2014; Aker 2015).
[20].The established cutoff for the presence of depression on the CESD scale is a score of 16.

However, we find no treatment effects on trust (measured by the WVS question that asks how much others can be trusted, MDE 0.14 std. dev.; all MDEs for psychological well-being in Online Appendix Table 22), locus of control (measured by Rotter's Locus of Control scale, Rotter 1966, MDE 0.14 std. dev.), and self-esteem (measured by Rosenberg's self-esteem scale, Rosenberg 1965, MDE 0.15 std. dev.). These results suggest that cash transfers may improve some aspects of psychological well-being but not others. This finding may partly account for contrasting evidence regarding the effect of windfall gains on psychological well-being; while the literature has documented many positive impacts (Kling, Liebman, and Katz 2007; Ssewamala, Han, and Neilands 2009; Lund et al. 2011; Ozer et al. 2011; Rosero and Oosterbeek 2011; Devoto et al. 2012; Finkelstein et al. 2012; Baird, De Hoop, and Özler 2013; Bandiera et al. 2013; Ghosal et al. 2013; Mendolia 2014; Banerjee et al. 2015a; Cesarini et al. 2016), some studies find little effect (Paxson and Schady 2010; Kuhn et al. 2011). Currently, it is difficult to discern a pattern—Paxson and Schady (2010) and Kuhn et al. (2011) studied depression and happiness, respectively, and found no effects of transfers on these outcomes, but these outcomes show positive responses to windfalls in both the present article and some previous ones (Ozer et al. 2011; Devoto et al. 2012; Finkelstein et al. 2012; Ssewamala et al. 2012; Banerjee et al. 2015a). Future work will have to establish which interventions affect which outcome measures.

In addition, we do not find an average effect of treatment on cortisol levels, either when measured as raw levels ("log cortisol" is the natural logarithm of the average of the two cortisol samples collected from each respondent) or when measured with controls (i.e., as residuals of an OLS regression of the log-transformed cortisol levels on dummies for having ingested food, drinks, alcohol, nicotine, or medications in the two hours preceding the interview, for having performed vigorous physical activity on the day of the interview, and for the time elapsed since waking, rounded to the next full hour). In both cases, the point estimates are small and not significant. We were reasonably powered to detect changes; the MDE for raw cortisol levels was 0.13 log units (that is, 5% of the control group mean). The MDE for cortisol levels with controls was also 0.13 log units. The null finding on cortisol contrasts with that of Fernald and Gunnar (2009), who show that children whose mothers participated in the Oportunidades program had lower cortisol levels. In comparison to self-reported measures, our cortisol results suggest that either cortisol is noisier and more difficult to affect with interventions than self-reported measures, or that the self-reports may be subject to experimental demand effects. The fact that we observe significant positive effects on some of these variables but not others argues against the demand effect explanation; if people were telling the surveyors "what they wanted to hear," we would have expected positive effects across the board. A further reason to think that noise plays a role is the fact that the treatment effects on self-report measures do not always have the same sign as those on cortisol; for instance, monthly transfer recipients have lower levels of depression than do lump-sum recipients, but higher levels of cortisol. Such discrepancies across variables are not unique to cortisol, but they illustrate a difficulty in using it as an outcome variable. In the absence of a gold standard measure that assesses well-being directly, it is not clear which variable should be given priority.

2.   Treatment Arms—In the following we discuss the differences in psychological well-being across treatment arms in more detail, with particular attention to cortisol. The corresponding results are reported in columns (3)–(5) of Table IV. First, overall psychological well-being is 0.14 std. dev. higher in female compared to male recipient households; this difference is significant at the 10% level and is mainly driven by self-esteem and cortisol. The MDEs for these comparisons ranged from 0.20 to 0.25 std. dev. The magnitude of the cortisol effect is 0.17 log units, which corresponds to a difference between female and male recipient households of 2.17 nmol/l. It is important to note that the individual comparisons of female and male recipient households to the spillover group are not significant, and male recipient households actually show a small and nonsignificant increase in cortisol levels (Online Appendix Tables 118, 127, 134). Nevertheless, the difference between the two groups is large when compared with the average difference in morning cortisol levels between depressed and healthy individuals reported in the literature (2.58 nmol/l; Knorr et al. 2010). This finding is particularly remarkable in light of the fact that we observe no other significant differences between male and female recipient households on any of our index variables; the only differences observed are in psychological well-being, with cortisol a main driver of the effect (as described already, we also observe lower levels of worries and higher levels of self-esteem in female recipient households). One possible explanation for these findings lies in the fact that (i) psychological well-being correlates highly with female empowerment in the cross section (Online Appendix Table 30), and that (ii) female empowerment shows a relatively large difference (although significant only at the 10% level) of 0.17 std. dev. between male and female recipient households (in fact, this difference is of roughly the same magnitude as the difference in psychological well-being between male and female recipient households, 0.14 std. dev.). Together, these findings suggest that the differential cortisol levels and other indicators of psychological well-being between male and female recipient households may reflect the reduced stress from increases in female empowerment. Of course this mechanism is speculative, but it provides a hypothesis for future research. The fact that the difference in female empowerment between male and female recipient households is not itself significant suggests either that the cortisol effect additionally reflects other changes that are not captured in female empowerment, and/or that cortisol responds better to interventions than measures based on self-report (note, however, that this should then also apply to psychological well-being, where we do observe differences; thus, the former explanation is more plausible). An alternative explanation is that men are under less stress to "provide" for the family when their wives receive transfers. This hypothesis is supported by the fact that the decrease in cortisol levels in female recipient households is, surprisingly, driven by men's cortisol levels, which are significantly reduced relative to those of the spillover group when women receive transfers, but are not reduced when the men themselves receive transfers (Online Appendix Tables 121–122). Women do not show changes in cortisol levels regardless of which spouse receives transfers.

Second, we find no overall difference in psychological well-being for monthly compared to lump-sum transfers, although depression is marginally lower in monthly recipient households, and cortisol is higher. The MDEs for the comparison of monthly to lump-sum transfers ranged from 0.22 to 0.26 std. dev. In monthly recipient households, cortisol levels

are 0.17 log units (2.17 nmol/l) higher than in households that receive their transfers as a lump sum. As shown in Online Appendix Tables 119, 128, and 135, this effect stems from an increase in cortisol levels relative to baseline in the monthly transfer recipient households; there is no significant decrease relative to baseline in lump-sum recipient households. This finding is surprising for two reasons. First, the cortisol effect is not accompanied by other differences in psychological well-being, suggesting that cortisol may reflect outcomes that are not well captured in self-report measures. Second, stress is strongly related to controllability, homeostasis, and stability, and given that monthly transfers increased food security to a greater extent than did lump-sum transfers, and food security correlates well with psychological well-being in the cross section (Online Appendix Table 30), we might have expected cortisol to be lower in monthly recipient households. A potential explanation for this surprising finding may lie in the fact that, as we discuss in greater detail below, households in the monthly condition seem to have had difficulty in saving or investing the transfers, possibly in spite of better intentions; thus, it is possible that the increased cortisol levels in this condition reflect the stress arising from this failure.

Finally, we find that cortisol levels are 0.12 log units (1.76 nmol/l) lower in households that received large transfers than in households that received small transfers. Concomitantly, we observe large additional gains in psychological well-being for large transfers. The overall index of psychological well-being is a full 0.26 std. dev. higher for larges than for small transfers. Apart from cortisol, this effect is driven by a 1.22-point difference in depression scores between large and small transfer recipients, a 0.24 std. dev. difference in stress scores, and a 0.19 std. dev. difference in life satisfaction. The MDEs for these comparisons ranged 0.21 to 0.27 std. dev.

Together, these findings provide support for our ingoing hypothesis that poverty alleviation would lead to improvements in psychological well-being. The results are mixed. Not all self-reported variables show treatment effects, and while cortisol levels are different across treatment arms, we observe no average treatment effect on cortisol. However, treatment effects on cortisol levels are generally (but not always) in the same direction as those on self-reported outcomes. Together, our results suggest that cortisol and measures of psychological well-being are useful complements to traditional measures of economic welfare, and may in some cases reflect aspects of welfare that are not well captured by more traditional measures.

IV.D. Consumption

1. Overall Effects—Table V shows detailed results for expenditure variables. These results are a subset of those prespecified; full results are reported in Online Appendix Tables 63–69. Expenditure is reported only for flow expenses, not durables, which are reported below. Data are monthly and top-coded at 99%; the Online Appendix (Tables 70–76) presents robustness checks using logarithmic coding. MDEs are shown in Online Appendix Table 21.

Overall, we find a significant increase in the monthly flow of nondurable expenditure of $36 PPP relative to a control group mean of $158 PPP at endline (23%). With the exception of temptation goods (alcohol and tobacco), transfers increase expenditures in all categories,

including food, medical and education expenditure, and social events. In absolute terms, the largest increase in consumption is food ($19 PPP, 19%). Spending on protein (meat and fish) is increased substantially in percentage terms ($5 PPP, 39%), while spending on staples (cereals) is less strongly increased ($2 PPP, 10%). Spending on medical care, education, and social events (e.g., weddings, funerals, recreation) increases significantly in percentage terms, but from relatively lower levels: Monthly medical expenditures increase by $3 PPP (38%), education expenditures increase by $1 PPP (23%), and social expenditures increase by $2 PPP (56%). Together, these findings broadly complement those of other cash transfer programs, which also report increases in consumption and, in particular, food expenditure (Maluccio and Flores 2005; Schady and Rosero 2008; Angelucci and De Giorgi 2009; Maluccio 2010; Gertler, Martinez, and Rubio-Codina 2012; Macours, Schady, and Vakis 2012; Skoufias, Unar, and Cossio 2013; Cunha 2014). The significant effect on the food security index reflects these findings (Aker et al. 2016).

Interestingly, the treatment effects are negative and not significantly different from 0 for alcohol and tobacco. We note, however, that one significant concern with these findings is that because of lack of power, we cannot rule out moderately sized positive treatment effects with confidence. For alcohol our MDE was $2.78 PPP per month, which is a 44% increase relative to the control group mean of $6.38 PPP. Analogously for tobacco, the MDE was $0.61 PPP, or 40% of the control group mean. Future studies with greater power can potentially provide more definitive evidence on the treatment effect on these outcomes.

A further potential concern when asking respondents about their consumption of alcohol and tobacco is desirability bias. Respondents may have told the research team what they thought the surveyors wanted to hear, and this effect may have been differentially large in the treatment group. However, three considerations suggest that this bias, if it existed, is unlikely to have influenced our results substantially. First, the survey team was kept distinct from the intervention team and denied any association when asked (although it remains possible that at least some respondents nevertheless suspected a connection). Second, we note that other variables do not show a treatment effect, even though for these variables social desirability would bias the results in the direction of finding an effect where there is none. For instance, in the case of educational and health outcomes, we find very little impact, despite the fact that if respondents were motivated to appear in a good light to the survey team, they would have had an incentive to overstate the benefits of the program in terms of these outcomes. Finally, we used a list randomization questionnaire in the endline to complement the direct elicitation of alcohol and tobacco expenditure. In this method, respondents are not directly asked whether they consumed alcohol or tobacco, but are presented with a list of five common activities such as visiting friends or talking on the phone, and asked how many of these activities they performed in the preceding week. The respondents were divided into three groups. One group was presented with only this short list; a second group was presented with the short list and an extra item, consuming alcohol; and for a third group, the extra item was consuming tobacco. Comparing the means across the different groups allows us to estimate the proportion of respondents who consumed alcohol and tobacco, without any respondent having to explicitly state that they did so. Online Appendix Table 29 suggests not only that there was no treatment effect on alcohol and tobacco consumption when using this method, but additionally that the estimates of alcohol and tobacco

consumption obtained through the list method are very similar (and if anything, lower) than those obtained through direct elicitation. Note, however, that a concern with this method is that it is noisy, and the results are therefore imprecise.

Finally, another potential concern regarding the expenditure results is that treatment households may spend money on different things because they spend it in different places (such as the kiosk where they withdraw money). However, anecdotally we know that most households withdraw their transfers immediately; at endline, the average M-Pesa balance in the treatment group is $4 PPP, significantly but not meaningfully higher than in the control group ($1; Online Appendix Table 33). As a result, this factor is unlikely to be of importance for the composition of the consumption bundle.

2. Treatment Arms—Are the expenditure effects different for different types of wealth changes, that is, transfers to the female versus the male, monthly versus lump-sum transfers, or large versus small transfers? These comparisons are shown in columns (3) to (5) of Table V.

First, the comparison of female and male recipient households reveals few differences in expenditure. This result is surprising in light of a large literature suggesting that households may not be unitary (Thomas 1990; Browning and Chiappori 1998; Duflo and Udry 2004). In this literature, a common test of whether households are unitary is precisely the test of income pooling, which asks whether expenditure shares are different when money is received by husbands versus wives. Several papers have found this to be the case in observational data (Thomas 1990; Hoddinott and Haddad 1995; Doss 1996; Lundberg, Pollak, and Wales 1997; Duflo 2003; Aker et al. 2016); in particular, female income is associated with larger expenditures on food and children. However, in our setting, one possible reason for the lack of significant differences between female versus male recipient households is low power in these comparisons; the MDE for total expenditure was $29 PPP, corresponding to 18% of the control group mean, and that for food expenditure was $21 PPP, corresponding to 20% of the control group mean.

The comparison of monthly and lump-sum recipient households in column (4) shows that expenditures in monthly recipient households do not differ significantly from lump-sum recipient households; none of the individual coefficients are significant, and joint significance is at $p = .13$. The MDEs for this comparison ranged from 20% of the control group mean for food consumption to 85% of the control group mean for medical expenditure for children. The MDE for total nondurable expenditure is 19% of the control group mean.

Finally, in large transfer recipient households, monthly expenditure is $21 PPP higher than in small transfer recipient households; Online Appendix Table 69 shows that small transfers increased expenditure by $30 PPP relative to control, and large transfers by $51 PPP. Thus, the treatment effects have a ratio of 1.7 for large relative to small transfers, whereas that of the transfers themselves is 3.8 for large compared to small transfers. The marginal propensity to spend out of the transfer is therefore decreasing in transfer size. Indeed, none of the individual expenditure categories show differential effects for large transfers. Our

MDEs for this comparison ranged from 20% of the control group mean for food consumption to 74% of the control group mean for medical expenditure for children. The MDE for total nondurable expenditure is again 19% of the control group mean.

IV.E.   Assets and Business Activities

1.   Overall Effects—In the following section we assess the impact of transfers on assets and investment, and explore average returns on investment and the possibility of savings and credit constraints.

We begin by estimating the treatment effect on asset ownership in Table VI, Panel A measured by asking respondents for the present value of a number of common assets. The variables reported here are a subset of those prespecified; full results are reported in the Online Appendix (Tables 49–62). MDEs are shown in Online Appendix Table 23. The overall treatment effect on assets amounts to $302 PPP relative to a control group mean of $495 PPP (61%), and is mainly driven by investment in livestock and durables. Livestock holdings increase by $83 PPP, a 50% increase relative to the control group mean, and 12% of the average transfer. Similarly, the value of durable goods owned by treatment households increased by $53 PPP relative to a control group mean of $207 PPP (an increase of 25%, or 7% of the average transfer), primarily due to purchases of furniture (beds, chairs, tables, etc.). Reported cash savings balances doubled as a result of cash transfers but from low initial levels (baseline mean of $11 PPP).

One of the most visible impacts of the transfer is investment in metal roofs. Cash transfers increase the likelihood of having a metal roof by 24 percentage points relative to a control group mean of 16% at endline. Because the average cost of a metal roof is $669 PPP, this effect corresponds to a $161 PPP increase in the value of roofs owned by treatment households, which in turn corresponds to 23% of the average transfer.[21]

The variables reported above are for durable assets; to get a full picture of investments that may have financial returns, we additionally created an index of nondurable investment, consisting of the total spending on agricultural inputs (seed, fertilizer, water, hired labor, livestock feed, livestock medicine, etc.), enterprise expenses (wages, electricity, water, transport, inventory, and other inputs), education expenditures (school and college fees, books, uniforms), and savings. Results are reported in Online Appendix Tables 143–148. We find an increase in nondurable investment by $23 PPP relative to a control group mean of $40 PPP (59%); on a percentage basis, this increase is thus very similar in magnitude to the increase in durable asset holdings described above.

Are these investments productive? We turn to two sources of evidence to address this question. First, metal roofs provide an investment return to households by obviating the need to replace and repair their thatched roofs, which costs on average $101 PPP a year (estimated from a sample of on average three respondents in each control group village). Given a cost of $669 PPP for a metal roof, we estimate a simple annual return on this

---

[21.]Incidentally, the fact that a large number of transfer recipients chose to acquire a metal roof suggests that they understood that the program was transitory, because by acquiring a metal roof they disqualified themselves from it.

investment of 15% (assuming no depreciation of metal roofs; this assumption is reasonable, as most respondents were unable to put an upper bound on the durability of metal roofs).

Second, Table VI, Panel B presents impacts of cash transfers on agricultural and business activities. The variables reported here are a subset of those prespecified; full results are reported in the Online Appendix (Tables 77–92). MDEs are shown in Online Appendix Table 24. Total revenue increases by $16 PPP relative to a control group mean of $49 PPP (33%, significant at the 1% level). However, we note that costs also increased by $13 PPP (52%, significant at the 1% level). As a result, we observe no significant treatment effect on self-reported profits, with a nonsignificant point estimate of "$0.21 PPP. The MDE for this effect was $10 PPP, or 50% of the control group mean. There is little evidence that transfers change the primary source of income for recipient households; they are no more or less likely than control households to report farming, wage labor, or nonagricultural businesses as their primary sources of income. The MDEs for these comparisons ranged from 6–9 percentage points. Additional detail on labor outcomes is reported in Section IV.F.

The returns to investment we observe are lower than those found in studies of cash transfers targeted at business owners and other select groups (de Mel, McKenzie, and Woodruff 2008; De Mel, McKenzie, and Woodruff 2012; Blattman, Fiala, and Martinez 2014; Fafchamps and Quinn 2015; McKenzie 2015; but see Fafchamps et al. 2011, who find no positive effects of cash grants in Ghana), but suggest that cash transfers may have the potential to increase long-term consumption even in a broader sample of the population (Aizer et al. 2016; Bleakley and Ferrie 2013). Blattman, Jamison, and Sheridan (2015) find increases in investment, labor supply, and profits after unconditional cash transfers to street youth in Liberia, but these effects are short-lived.

2. Treatment Arms—We now briefly discuss differences in asset holdings and business outcomes across treatment arms. First, quite naturally, we find that large transfers increase asset holdings by an additional $279 PPP relative to small transfers. Second, we find little evidence that asset holdings and business outcomes are different when transfers are made to the woman rather than the man, except that female recipient households are 11 percentage points less likely to invest in metal roofs. However, the *MDEs* for these comparisons were large, ranging from 19% of the control group mean for durables to 129% for savings.

Finally, the randomization in transfer timing (lump-sum versus monthly) allows us to ask whether households are savings or credit constrained (Ashraf, Karlan, and Yin 2006; Dupas and Robinson 2013a). Specifically, the permanent income hypothesis (Modigliani and Brumberg 1954; Friedman 1957) predicts that households invest the balance of their transfers to smooth consumption over time. However, a significant portion of the transfer was spent on consumption goods. One possible explanation of this finding is that households face lumpy investment opportunities and are constrained in their ability to save and borrow; in this case, we would expect fewer purchases of expensive assets such as metal roofs among monthly transfer recipients, because the savings constraint would prevent this group from saving their transfer to buy the asset, and the credit constraint would prevent it from borrowing against the promise of future transfer. Indeed, in column (4) of Table VI we find that endline asset holdings of monthly recipient households are significantly lower than

those of lump-sum recipients (although note that this effect does not survive FWER adjustment in Table II). In particular, monthly recipients are 12 percentage points less likely to acquire a metal roof, and instead use more of the transfer for current consumption, evident in a significantly higher food security index (Table II). Thus, monthly recipient households may be both credit and savings constrained. The finding that lump-sum recipient households are more likely to make large investments mirrors that of Barrera-Osorio et al. (2008), who find that bundling the payments of a conditional cash transfer program at the time when children have to reenroll in school increases enrollment rates.

The fact that program participation required signing up for mobile money accounts, which are a low-cost savings technology (people could have chosen to accumulate their transfer in their M-Pesa account and even add additional funds), suggests that the savings constraint at work is more social or behavioral than due purely to the lack of access to a savings technology. Anecdotally, we know that recipients were often asked by family members to share the transfer. In the case of monthly transfers, a single installment of which usually cannot be used to buy large, lumpy assets, these requests may be harder to refuse, while lump-sum transfer recipients might have an easier time arguing that the entire transfer is needed to pay for the planned purchase. An additional factor may be that households had more time to plan what to do with the lump-sum transfers, since they arrived half-way through the treatment period on average, while monthly transfers began soon after the announcement for all households.

The results on cortisol levels (Table IV) provide a tantalizing complement to this interpretation. We find that cortisol levels are significantly higher for households who receive monthly transfers than for those who receive lump-sum transfers—and in fact, significantly elevated even relative to cortisol levels for the control group (see Online Appendix Tables 119, 128, and 135). This is despite the fact that immediate pressures on the lives of these recipients have decreased, as evident, for example, in the significant increase in food security. It is conceivable that the increase in cortisol levels for monthly transfer recipients reflects the stress associated with having to decide continually how to spend the transfers or an inability to save, whereas the transfers of lump-sum recipients are safely invested in metal roofs.

IV.F.   Do Cash Transfers Create Dependency?

An important policy question is whether cash transfers create dependency. We address this question in two ways. First, we study the temporal evolution of the treatment effect. This variation in the present study is not sufficient to obtain reliable estimates for the evolution of the treatment effect over time; however, in Online Appendix Figure 4 and Table 20, we show treatment effect estimates for households receiving their last transfer less than one month ago, households receiving their last transfer one to four months ago, and households receiving their last transfer four or more months ago, and we do this separately for monthly and lump-sum recipient households. The observed average impact on asset holdings, consumption, and food security persists over time, although it is larger for more recent transfers. For agricultural and business revenue, psychological well-being, health, education, and female empowerment, we find no strong indication of changing impacts over time;

however, the treatment effects in the individual time bins are small overall and not distinguishable from 0 in this restricted and highly disaggregated sample. The increase in revenue at the longest time horizon is not individually significant, but the MDE is 46% of the control group mean. Psychological well-being is not significantly elevated at the longest time horizon, despite a relatively small MDE of 0.22 std. dev.

Second, we ask whether transfers reduce labor supply. In Table VI, we find no effects of transfer receipt on dummies for wage labor versus the own farm being the primary source of income. Online Appendix Tables 137–142 show additional variables related to labor provision, and find no effects of the transfers on the proportion of working-age household members who spent any time in the preceding 12 months doing casual labor (MDE 6 percentage points) or working in a salaried job (MDE 3 percentage points), the likelihood of "casual labor" or "salaried job" being ranked among the top three sources of income for the household (MDE 6 percentage points), or the amount spent on hiring labor for agricultural activities (MDE $6 PPP, control group mean $0 PPP). We do, however, find a significantly positive effect on the number of income-generating activities reported by the household, suggesting that households diversified their income-generating activities. Thus, it does not appear that cash transfers affect labor supply negatively, in line with existing findings (Posel, Fairburn, and Lund 2006; Ardington, Case, and Hosegood 2009; Banerjee et al. 2015b), and that they may in fact affect it positively.[22] We did not study how transfers affect labor supply by children (Edmonds 2006; Edmonds and Schady 2012).

V. Conclusion

This article reports the results from an impact evaluation of UCTs in a sample of poor households in western Kenya. The study differs from previous UCT experiments in that the transfers were entirely unconditional and were relatively large and concentrated in time, with randomization of recipient gender, transfer timing, and transfer magnitude. In addition, we observe a large number of outcome variables and pay particular attention to psychological well-being and cortisol levels.

We find that treatment households increased both consumption and savings (in the form of durable good purchases and investment in their self-employment activities). In particular, we observe increases in food expenditures and food security, but not spending on temptation goods. Households invest in livestock and durable assets (notably metal roofs), and we show that these investments lead to increases in revenue from agricultural and business activities, although we find no significant effect on profits at this short time horizon. We also observe no evidence of conflict resulting from the transfers; on the contrary, we report large increases in psychological well-being and an increase in female empowerment with a large spillover effect on nonrecipient households in treatment villages. Thus, these findings suggest that simple cash transfers may not have the perverse effects that some policy makers feel they would have, which has led to a clear policy preference for in-kind or skills transfers (Bandiera et al. 2013; Banerjee et al. 2015a; Brune et al. 2015) and conditional transfers

---

[22].This lack of an effect of windfalls on labor supply may not hold in developed countries; for instance, Cesarini et al. (2015) report a decrease in labor supply in Swedish households after a cash windfall.

(Sadoulet, Janvry, and Davis 2001; Maluccio and Flores 2005; Attanasio and Mesnard 2006; Ferreira, Filmer, and Schady 2009; Filmer and Schady 2009; Maluccio 2010; Maluccio, Murphy, and Regalia 2010; Baird, McIntosh, and Özler 2011; Baird et al. 2012; Gertler, Martinez, and Rubio-Codina 2012; Macours, Schady, and Vakis 2012; Akresh, de Walque, and Kazianga 2013; Baird, De Hoop, and Özler 2013; Barham, Macours, and Maluccio 2013; Skoufias, Unar, and Cossio 2013).

The three treatment arms included in this study—transfers to the woman versus the man in the household, monthly versus lump-sum transfers, and large versus small transfers—enable us to speculate about the likely impact of varying these design features in existing cash transfer programs. For large transfers, the results are relatively unambiguous: they produce more desirable results on most outcome measures, including asset holdings, consumption, food security, psychological well-being, and female empowerment (although not revenue, health, and education), but the returns to transfer size appear to be decreasing. Monthly transfers are superior to lump-sum transfers in terms of their effects on food security, whereas lump-sum transfers show larger effects than monthly transfers on asset holdings. Finally, making transfers to the woman in the household produces weakly larger treatment effects than transfers to the male on female empowerment and psychological well-being. Together, these results suggest that when policy makers consider the welfare implications of different design choices for UCTs, they may come to different conclusions depending on how they weight different dimensions of welfare relative to one another. However, an important caveat to our findings is low statistical power in the cross-randomizations of recipient gender and transfer timing and magnitude; null effects should therefore not be overinterpreted. We provide a guide to the interpretation of such null results by reporting MDEs. A further limitation is the relatively short time horizon of our endline. We stress, however, that our results can provide useful evidence on the short-term impact of cash transfers and on the existence of savings and credit constraints. In addition, this initial analysis provides useful policy guidance for governments and other entities which redistribute in ways other than cash because of concerns about how the money will be spent (e.g., on temptation goods).

The treatment effects on levels of the stress hormone cortisol raise a number of intriguing questions for future research. The finding that cortisol levels are significantly lower when transfers are made to the wife rather than the husband is surprising, because it occurs in the absence of differential effects between male and female recipients on other outcomes. This result therefore raises the question of whether cortisol may track particular aspects of welfare with greater sensitivity than traditional measures; for instance, they may reflect differences in female empowerment that are not visible in self-report measures such as our female empowerment index. More generally, our results suggest that cortisol levels can respond to economic interventions, and given its other advantages (objectivity, correlation with psychological well-being, implications for long-term health, and ease of collection), that it may be a useful complement to existing outcome measures in impact evaluation.

The present findings raise a number of questions for future research. First, as mentioned, the long-term effects of UCTs are topic of crucial importance for both economists and policy makers and are still incompletely understood, especially in developing countries (Bleakley

and Ferrie 2013; Aizer et al. 2016). Second, our study was not well powered to study the equilibrium effects of cash transfers at the level of the local economy: whether UCTs lead to changes in prices, wages, and crime at the local level remains an important topic for future investigation. Third, although we study the effect of UCTs on food expenditure, we do not address their effects on calorie consumption; this also remains a topic for future work (Deaton and Subramanian 1996). Fourth, the large spillover effects on female empowerment we report here deserve further investigation. Fifth, we do not study heterogeneous treatment effects of UCTs in detail; future work might investigate if they work differentially well for different target groups. Finally, from a policy point of view, the present study is only a small step in that adds to the growing body of evidence showing that UCTs have broadly "positive" welfare impacts, with little evidence for "negative" effects such as increases in conflict or temptation good consumption. This is encouraging, but it is only a starting point: what is needed now, in our view, are studies that compare the effect of cash transfers to those of other interventions that have been shown to be effective in improving outcomes in developing countries. For instance, are UCTs more or less effective than ultrarpoor graduation programs such as that studied by Banerjee et al. (2015a), and on what dimensions? Are UCTs delivered to a population simply chosen for being poor, such as in this study, more or less effective than transfers directed at recipients chosen in other ways, for example, caretakers of orphans, pensioners, or business owners (De Mel, McKenzie, and Woodruff 2008; Blattman, Fiala, and Martinez 2014; Fafchamps and Quinn 2015; McKenzie 2015)? In addition to revealing which interventions are most effective in achieving specific policy goals, such studies would facilitate the interpretation of results across contexts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Appendix

APPENDIX TABLE A.1

Ex Post Minimum Detectable Effect Sizes (MDEs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Treatment effect		**Female recipient**		**Monthly transfer**		**Large transfer**	
	Control mean	**MDE**	**Percent of control mean**	**MDE**	**Percent of control mean**	**MDE**	**Percent of control mean**	**MDE**	**Percent of control mean**
Value of nonland assets (US$)	494.80 (415.32)	76.29	0.15	141.07	0.29	128.57	0.26	137.46	0.28
Nondurable expenditure (US$)	157.61 (82.18)	16.39	0.10	28.77	0.18	29.98	0.19	29.36	0.19
Total revenue, monthly (US$)	48.98 (90.52)	16.46	0.34	29.71	0.61	31.01	0.63	24.84	0.51
Food security index	0.00 (1.00)	0.17		0.25		0.30		0.28	
Health index	0.00 (1.00)	0.17		0.24		0.28		0.25	
Education index	0.00 (1.00)	0.16		0.25		0.27		0.24	
Psychological well-being index	0.00 (1.00)	0.14		0.21		0.24		0.22	
Female empowerment index	0.00 (1.00)	0.20		0.29		0.33		0.30	

Notes. Ex post power calculations and MDEs for main outcome indexes and treatment arms. Outcome variables are listed on the left. The unit of observation is the household for all variables expect psychological well-being, where it is the individual. The sample includes all households and individuals, except for the intrahousehold index, where it is restricted to cohabitating couples, and for the education index, where it is restricted to households with school-age children. For each outcome variable, we report the control group mean and standard deviation in column (1). In columns (2), (4), (6), and (8), we report the MDEs for the main treatment effect and the comparison between treatment arms, respectively, calculated ex post using a significance level of 0.05 and power of 80%. In columns (3), (5), (7), and (9), we report, for monetary outcome variables, the MDE as a proportion of the control group mean for the main treatment effect and the treatment arms, respectively.

*
denotes significance at 10%

**
5%, and

***
1% levels.

APPENDIX TABLE A.2

Outcome Variable Descriptions

Index variables	
Value of nonland assets	Total value (in 2012 US$ PPP) of all nonland assets owned by the household, including savings, livestock, durable goods, and metal roofs.
Nondurable expenditure	Total monthly spending (in 2012 US$ PPP) on nondurables, including food, temptation goods, medical care, education expenditures, and social expenditures.
Total revenue, monthly	Total monthly revenue (in 2012 US$ PPP) from all household enterprises, including revenue from agriculture, stock and flow revenue from animals owned by the household, and revenue from all nonfarm enterprises owned by any household member.
Food security index	A standardized weighted average of the (negatively coded) number of times household adults and children skipped meals, went whole days without food, had to eat cheaper or less preferred food, had to rely on others for food, had to purchase food on credit, had to hunt for or gather food, had to beg for food, or went to sleep hungry in the preceding week; a

Index variables	
	(negatively coded) indicator for whether the respondent went to sleep hungry in the preceding week; the (positively coded) number of times household members ate meat or fish in the preceding week; (positively coded) indicators for whether household members ate at least two meals per day, ate until content, had enough food for the next day, and whether the respondent ate protein in the last 24hours; and the (positively coded) proportion of household members who ate protein in the last 24 hours, and proportion of children who ate protein in the last 24 hours.
Health index	A standardized weighted average of the (negatively coded) proportion of household adults who were sick or injured in the last month, the (negatively coded) proportion of household children who were sick or injured in the last month, the (positively coded) proportion of sick or injured family members for whom the household could afford treatment, the (positively coded) proportion of illnesses for which a doctor was consulted, the (positively coded) proportion of newborns who were vaccinated, the (positively coded) proportion of children below age 14 who received a health checkup in the preceding six months, the (negatively coded) proportion of children under age 5 who died in the preceding year, and a children's anthropometrics index consisting of BMI, height-for-age, weight-for-age, and upper-arm circumference relative to WHO development benchmarks.
Education index	A standardized weighted average of the proportion of household children enrolled in school and the amount spent by the household on educational expenses per child.
Psychological well-being index	A standardized weighted average of their (negatively coded) scores on the CESD scale (Radloff 1977), a custom worries questionnaire (negatively coded), Cohen's stress scale (Cohen, Kamarck, and Mermelstein 1983) (negatively coded), their response to the World Values Survey happiness and life satisfaction questions, and their log cortisol levels adjusted for confounders (negatively coded).
Female empowerment index	A standardized weighted average of a measure of two other indexes, a violence and an attitude index. The violence index is a weighted standardized average of the frequency with which the respondent reports having been physically, sexually, or emotionally abused by her husband in the preceding six months; the attitude index is a weighted standardized average of a measure of the respondent's view of the justifiability of violence against women, and a scale of male-focused attitudes.
Psychological well-being	
Log cortisol (no controls)	Log of salivary cortisol levels in nmol/L, taken as the average of level across two samples.
Log cortisol (with controls)	Residuals of an OLS regression of the log-transformed cortisol levels on dummies for having ingested food, drinks, alcohol, nicotine, or medications in the two hours preceding the interview, for having performed vigorous physical activity on the day of the interview, and for the time elapsed since waking, rounded to the next full hour
Depression (CESD)	The standardized total of the score from the 20 elements of the CESD questionnaire (Radloff 1977).
Worries	The standardized total of the score from the 16 elements of a custom worries questionnaire.
Stress (Cohen)	The standardized total of score from four elements of Cohen's stress scale (Cohen, Kamarck, and Mermelstein 1983).
Happiness (WVS)	The standardized numerical response to the World Values Survey happiness question: Taking all things together, would you say you are "very happy" (1), "quite happy" (2), "not very happy" (3), or "not at all happy" (4)?
Life satisfaction (WVS)	The standardized numerical response to the World Values Survey life satisfaction question: All things considered, how satisfied are you with your life as a whole these days on a scale of 1 to 10? (1 = very dissatisfied … 10 = very satisfied)
Trust (WVS)	The standardized numerical response to the World Values Survey life satisfaction question: generally speaking, would you say that most people can be trusted (1) or that you need to be very careful in dealing with people (2)?
Locus of control	The standardized weighted average of the total score on Rotter's locus of control questionnaire and the numerical response the World Values Survey locus of control question (Some people believe that individuals can decide their own destiny, while others think that it is impossible to escape a predetermined fate. Please tell me which comes closest to your view on this scale on which 1 means "everything in life is determined by fate" and 10 means "people shape their fate themselves.") Higher scores indicate a more internal locus of control.
Optimism (Scheier)	The standardized total score on the 6-question Scheier optimism questionnaire.
Self-esteem (Rosenberg)	The standardized total score on the 10-question Rosenberg optimism questionnaire.

Index variables	
Consumption	
Food total	The combined monthly total (in 2012 US$ PPP) of all spending on food by the household and the value of all food produced from agriculture and livestock that was consumed by the household (calculated as the monthly average of total production in the last year).
Cereals	The monthly total spending (in 2012 US$ PPP) by all members of the household on all cereal grains, flours, breads, pastas, cakes, and biscuits.
Meat & fish	The monthly total spending (in 2012 US$ PPP) by all members of the household on all meat and fish.
Alcohol	The monthly total spending (in 2012 US$ PPP) by all members of the household on alcoholic beverages.
Tobacco	The monthly total spending (in 2012 US$ PPP) by all members of the household on tobacco products, including cigarettes, cigars, tobacco, snuff, khatt, and miraa.
Social expenditure	The monthly total spending (in 2012 US$ PPP) by all members of the household on ceremonies, weddings, funerals, dowries/bride prices, charitable donations, village elder fees, and recreation or entertainment.
Medical expenditure past month	The monthly total spending (in 2012 US$ PPP) on medical care for all household members including consultation fees, medicines, hospital costs, lab test costs, ambulance costs, and related transport.
Education expenditure	The monthly total spending (in 2012 US$ PPP) for household members on school/college fees, uniforms, books, and other supplies.
Assets	
Value of livestock	The total value (in 2012 US$ PPP) of all cattle (cows, bulls, or calves), birds (chicken, turkeys, doves, quails, etc.), and small livestock (pigs, sheep, goats, etc.) owned by the household.
Value of durable goods	The total value (in 2012 US$ PPP) of all durable goods, including transportation, furniture, agricultural equipment, appliances, radios and televisions, and phones owned by the household.
Value of savings	The total value (in 2012 US$ PPP) of all savings held by all members of the household in any account (post bank, SACCO, village bank, M-Pesa, Zap, ROSCA, commercial bank, microfinance institution, etc.), at home, or held by friends, relatives, or colleagues.
Land owned (acres)	The total value (in 2012 US$ PPP) of all land owned by any household member.
Has nonthatched roof (dummy)	An indicator variable taking the value of 1 if the household's primary residence has a metal or concrete roof and 0 otherwise.
Business activities	
Wage labor primary income (dummy)	An indicator variable taking the value of 1 if the household's primary source of income is wage labor and 0 otherwise.
Own farm primary income (dummy)	An indicator variable taking the value of 1 if the household's primary source of income is a farm owned by the household and 0 otherwise.
Nonagricultural business owner (dummy)	An indicator variable taking the value of 1 if any member of the household is full or part owner of a nonagricultural enterprise and 0 otherwise.
Total revenue, monthly	Total monthly revenue (in 2012 US$ PPP) from all household enterprises, including revenue from agriculture, stock and flow revenue from animals owned by the household, and revenue from all nonfarm enterprises owned by any household member.
Total expenses, monthly	Total in 2012 US$ PPP of all expenses on agricultural enterprises including seeds, fertilizers/herbicides/pesticides, hired machines, water, and labor, and all expenses on nonagricultural enterprises including machinery/durable goods, inputs/inventory, salaries/wages, transport, electricity, and water.
Total profit, monthly	Total imputed profit in 2012 US$ PPP from all agricultural and nonagricultural enterprises owned by the household.

APPENDIX TABLE A.3

Preanalysis Plan Discrepancies

Preanalysis pan	Modification	Location
Outcome variables		
Total assets	Include roof values and omit land values	Main paper Tables I–II
Nondurable expenditures	Omit durable spending and spending on home repair (creates a better measure of flow consumption)	Main paper Tables I–II
Female empowerment (index)	Replace dummies for whether violent incident was reported with variables for the number of types of incidents reported (more granular measure)	Main paper Tables I–II
Financial variables	Omit value of outstanding loans and inability to pay loans (not relevant for remittance results)	OA Table 34
Preference & political variables	Omit for space reasons	Omitted
Village-level price index	Substitute a weighted average of all individual prices for a sum	OA Table 149
Within-village treatment effect estimation (PAP equations 3–6)	In addition to conditioning on baseline outcomes, we include a missing-value indicator variable	Main paper equations 5–8
Heterogenous treatment effects effects (PAP equations 8–12)	Omit for space reasons	Omitted
Temporal dynamics: early transfers versus late transfers (PAP equations 13–17)	Omit for space reasons	Omitted
Temporal dynamics: treatment effects broken out for transfers completed <1 month, 1–3 months, 4–6 months, and 6–12 months before endline (PAP equation 18).	Changed timing buckets to < 1 month, 1–4 months, and >4 months before endline (correction of PAP typo and increase of power at long time horizons)	OA Table 20 and Figure 4
Endline timing: create separate treatment dummies based on the date of the endline survey (PAP section 5.5)	Replaced this analysis with the original treatment effect calculation with controls for endline timing	OA Tables 16–19
Midline treatment effects	Omitted for space reasons	Omitted
Village-level general equilibrium effects (PAP equation 20)	Instead of regressing at the individual level, we collapse to the village level	OA Tables 149–159

References

Aggarwal Bharat B., Shishodia Shishir, Sandur Santosh K., Pandey Manoj K., and Sethi Gautam, "Inflammation and Cancer: How Hot Is the Link?," Biochemical Pharmacology, 72 (2006), 1605–1621. [PubMed: 16889756]

Aizer Anna, Eli Shari, Ferrie Joseph, and Lleras-Muney Adriana, "The Long-Run Impact of Cash Transfers to Poor Families," American Economic Review, 106 (2016), 935–971. [PubMed: 28713169]

Aker Jenny C., "Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo," World Bank Economic Review (2015). doi:10.1093/wber/lhv055.

Aker Jenny C., Boumnijel Rachid, McClelland Amanda, and Tierney Niall, "Payment Mechanisms and Anti-Poverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger," Economic Development and Cultural Change, 65 (2016), 1–37.

Akresh Richard, de Walque Damien, and Kazianga Harounan, "Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality," Policy Research Working Paper 6340, The World Bank, 2013.

Anderson Michael, "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 (2008), 1481–1495.

Angelucci Manuela, and De Giorgi Giacomo, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," American Economic Review, 99 (2009), 486–508.

Angrist Joshua D., Imbens Guido W., and Rubin Donald B., "Identification of Causal Effects Using Instrumental Variables," Journal of the American Statistical Association, 91 (1996), 444–455.

Ardington Cally, Case Anne, and Hosegood Victoria, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," American Economic Journal: Applied Economics, 1 (2009), 22–48. [PubMed: 19750139]

Ashraf Nava, Karlan Dean, and Yin Wesley, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," Quarterly Journal of Economics, 121 (2006), 635–672.

Attanasio Orazio, and Mesnard Alice, "The Impact of a Conditional Cash Transfer Programme on Consumption in Colombia," Fiscal Studies, 27 (2006), 421–442.

Baird Sarah J., De Hoop Jacobus, and Özler Berk, "Income Shocks and Adolescent Mental Health," Journal of Human Resources, 48 (2013), 370–403.

Baird Sarah J., Garfein Richard S., Craig T. McIntosh, and Özler Berk, "Effect of a Cash Transfer Programme for Schooling on Prevalence of HIV and Herpes Simplex Type 2 in Malawi: A Cluster Randomised Trial," Lancet, 379 (2012), 1320–1329. [PubMed: 22341825]

Baird Sarah J., McIntosh Craig, and Özler Berk, "Cash or Condition? Evidence from a Cash Transfer Experiment," Quarterly Journal of Economics, 126 (2011), 1709–1753.

Baird Sarah J., and Özler Berk, "Examining the Reliability of Self-Reported Data on School Participation," Journal of Development Economics, 98 (2012), 89–93.

Bandiera Oriana, Burgess Robin, Das Narayan, Gulesci Selim, Rasul Imran, and Sulaiman Munshi, "Can Basic Entrepreneurship Transform the Economic Lives of the Poor?," SSRN Scholarly Paper 2266813, 2013.

Banerjee Abhijit V., Duflo Esther, Goldberg Nathanael, Karlan Dean, Osei Robert, Pariente William, Shapiro Jeremy, Thuysbaert Bram, and Udry Christopher, "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," Science, 348 (2015a). doi:10.1126/science.1260799.

Banerjee Abhijit V., Hanna Rema, Kreindler Gabriel, and Olken Benjamin A, "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs Worldwide," HKS Working Paper, 2015b.

Barham Tania, Macours Karen, and Maluccio John A., "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, 2013.

Barrera-Osorio Felipe, Bertrand Marianne, Linden Leigh, and Perez-Calle Francisco, "Conditional Cash Transfers in Education Design Features, Peer and Sibling Effects Evidence from a Randomized Experiment in Colombia," NBER Working Paper 13890, 2008.

Benhassine Najy, Devoto Florencia, Duflo Esther, Dupas Pascaline, and Pouliquen Victor, "Turning a Shove into a Nudge? A Labeled Cash Transfer for Education," NBER Working Paper 19227, 2013.

Blattman Christopher, Fiala Nathan, and Martinez Sebastian, "Generating Skilled Employment in Developing Countries: Experimental Evidence from Uganda," Quarterly Journal of Economics, 129 (2014), 697–752.

Blattman Christopher, Jamison Julian, and Sheridan Margaret, "Reducing Crime and Violence: Experimental Evidence on Adult Noncognitive Investments in Liberia," NBER Working Paper 21204, 2015.

Bleakley Hoyt, and Ferrie Joseph P., "Up from Poverty? The 1832 Cherokee Land Lottery and the Long-Run Distribution of Wealth," NBER Working Paper 19175.

Bobonis Gustavo J., Gonzalez-Brenes Melissa, and Castro Roberto, "Public Transfers and Domestic Violence: The Roles of Private Information and Spousal Control," American Economic Journal: Economic Policy, 5 (2013), 179–205.

Browning Martin, and Chiappori Pierre, "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests," Econometrica, 66 (1998), 1241–1278.

Brune Lasse, Giné Xavier, Goldberg Jessica, and Yang Dean, "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi," Economic Development and Cultural Change, 64 (2015), 187–220.

Burbidge John B., Magee Lonnie, and Robb A. Leslie, "Alternative Transformations to Handle Extreme Values of the Dependent Variable," Journal of the American Statistical Association, 83 (1988), 123–127.

Cameron A. Colin, Gelbach Jonah B., and Miller Douglas L., "Robust Inference with Multiway Clustering," Journal of Business & Economic Statistics, 29 (2011), 238–249.

Casey Katherine, Glennerster Rachel, and Miguel Edward, "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan," Quarterly Journal of Economics, 127 (2012), 1755–1812.

Cesarini David, Lindqvist Erik, Notowidigdo Matthew J., and Östling Robert, "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries," NBER Working Paper 21762, 2015.

Cesarini David, Lindqvist Erik, Ostling Robert, and Wallace Bjorn, "Estimating the Causal Impact Of Wealth on Health: Evidence from the Swedish Lottery Players," Quarterly Journal of Economics (2016). doi:10.1093/qje/qjw001.

Chang Andrew C., and Li Phillip. 2015 "Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say 'Usually Not'," Finance and Economics Discussion Series 2015–083, 2015. doi:10.17016/FEDS.2015.083.

Chemin Matthieu, Haushofer Johannes, and Jang Chaning, "Health Insurance Reduces Stress: Evidence from a Randomized Experiment in Kenya," 2016, available at http://econ.as.nyu.edu/docs/IO/42071/Chemin_Haushofer_Jang_Insurance_2016.03.10.pdf.

Cohen Sheldon, Kamarck Tom, and Mermelstein Robin, "A Global Measure of Perceived Stress," Journal of Health and Social Behavior, 24 (1983), 385–396. [PubMed: 6668417]

Coussens Lisa M., and Werb Zena, "Inflammation and Cancer," Nature, 420 (2002), 860–867. [PubMed: 12490959]

Cunha Jesse M., "Testing Paternalism: Cash versus In-Kind Transfers," American Economic Journal: Applied Economics, 6 (2014), 195–230.

Currie Janet, and Gahvari Firouz, "Transfers in Cash and in Kind: Theory Meets the Data," Journal of Economic Literature, 46 (2008), 333–383.

Deaton Angus, and Subramanian Shankar, "The Demand for Food and Calories," Journal of Political Economy, 104 (1996), 133–162.

de Chaisemartin Clement, "Defying the LATE? Identification of Local Treatment Effects when the Instrument Violates Monotonicity," Quantitative Economics, forthcoming.

De Mel Suresh, McKenzie David, and Woodruff Christopher, "Returns to Capital in Microenterprises: Evidence from a Field Experiment," Quarterly Journal of Economics, 123 (2008), 1329–1372.

———, "One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka," Science, 335 (2012), 962–966. [PubMed: 22363007]

Devoto Florencia, Duflo Esther, Dupas Pascaline, Parient William, and Pons Vincent, "Happiness on Tap: Piped Water Adoption in Urban Morocco," American Economic Journal: Economic Policy, 4 (2012), 68–99.

Doss Cheryl R., "Testing among Models of Intrahousehold Resource Allocation," World Development, 24 (1996), 1597–1609.

Duflo Esther, "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa," World Bank Economic Review, 17 (2003), 1–25.

Duflo Esther, Glennerster Rachel, and Kremer Michael, "Using Randomization in Development Economics Research: A Toolkit," in Handbook of Development Economics, Vol. 4, edited by Schultz T. Paul and Strauss John A. (Amsterdam: Elsevier, 2007).

Duflo Esther, and Udry Christopher, "Intrahousehold Resource Allocation in Cote D'Ivoire: Social Norms, Separate Accounts and Consumption Choices," NBER Working Paper 10498, 2004.

Dupas Pascaline, and Robinson Jonathan, "Savings Constraints and Microenterprise: Evidence from a Field Experiment in Kenya," American Economic Journal: Applied Economics, 5 (2013a), 163–192.

———, "Why Don't the Poor Save More? Evidence from Health Savings Experiments," American Economic Review, 103 (2013b), 1138–1171. [PubMed: 29533047]

Edmonds Eric V., "Child Labor and Schooling Responses to Anticipated Income in South Africa," Journal of Development Economics, 81 (2006), 386–414.

Edmonds Eric V., and Schady Norbert, "Poverty Alleviation and Child Labor," American Economic Journal: Economic Policy, 4 (2012), 100–124.

Efron Bradley, and Tibshirani Robert, An Introduction to the Bootstrap (Boca Raton, FL: CRC Press, 1993).

Fafchamps Marcel, McKenzie David, Quinn Simon R., and Woodruff Christopher, "When is Capital Enough to Get Female Microenterprises Growing? Evidence from a Randomized Experiment in Ghana," NBER Working Paper 17207, 2011.

Fafchamps Marcel, and Quinn Simon, "Aspire," Journal of Development Economics, forthcoming.

Fernald Lia, and Gunnar Megan R., "Effects of a Poverty-Alleviation Intervention on Salivary Cortisol in Very Low-Income Children," Social Science & Medicine, 68 (2009), 2180–2189. [PubMed: 19406546]

Ferracuti Stefano, Seri Stefano, Mattia Donatella, and Cruccu Giorgio, "Quantitative EEG Modifications during the Cold Water Pressor Test: Hemispheric and Hand Differences," International Journal of Psychophysiology, 17 (1994), 261–268. [PubMed: 7806469]

Ferreira Francisco H. G., Filmer Deon, and Schady Norbert, "Own and Sibling Effects of Conditional Cash Transfer Programs: Theory and Evidence from Cambodia," Policy Research Working Paper 5001, World Bank, 2009.

Filmer Deon, and Schady Norbert, "Are There Diminishing Returns to Transfer Size in Conditional Cash Transfers?," Policy Research Working Paper 4999, World Bank, 2009.

Finkelstein Amy, Taubman Sarah, Wright Bill, Bernstein Mira, Gruber Jonathan, Newhouse Joseph P., Allen Heidi, Baicker Katherine, and the Oregon Health Study Group, "The Oregon Health Insurance Experiment: Evidence from the First Year," Quarterly Journal of Economics, 95 (2012), 1668–1679.

Friedman Milton, Theory of the Consumption Function (Princeton, NJ: Princeton University Press, 1957).

Gertler Paul, Martinez Sebastian, and Rubio-Codina Marta, "Investing Cash Transfers to Raise Long-Term Living Standards," American Economic Journal: Applied Economics, 4 (2012), 164–192.

Ghosal Sayantan, Jana Smarajit, Mani Anandi, Mitra Sandip, and Roy Sanchari, "Sex Workers, Stigma and Self-Belief: Evidence from a Psychological Training Program in India," Working Paper, University of Warwick, 2013.

Goldstein Jacob, "Is it Nuts to Give to the Poor without Strings Attached?," New York Times, August 13, 2013, available at http://www.nytimes.com/2013/08/18/magazine/is-it-nuts-to-give-to-the-poor-without-strings-attached.html?_r=0.

Hammen Constance, "Stress and Depression," Annual Review of Clinical Psychology, 1 (2005), 293–319.

Haushofer Johannes, and Fehr Ernst, "On the Psychology of Poverty," Science, 344 (2014), 862–867. [PubMed: 24855262]

Haushofer Johannes, Reisinger James, and Shapiro Jeremy, "Is Your Gain My Pain?: Negative Psychological Externalities of Wealth Changes," working paper, Princeton University, 2015.

Hidrobo Melissa, Peterman Amber, and Heise Lori, "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador," American Economic Journal: Applied Economics, 8 (2016), 284–303.

Hoddinott John, and Haddad Lawrence, "Does Female Income Share Influence Household Expenditures? Evidence from Côte d'Ivoire," Oxford Bulletin of Economics and Statistics, 57 (1995), 77–96.

Holsboer Florian, "The Corticosteroid Receptor Hypothesis of Depression," Neuropsychopharmacology, 23 (2000), 477–501. [PubMed: 11027914]
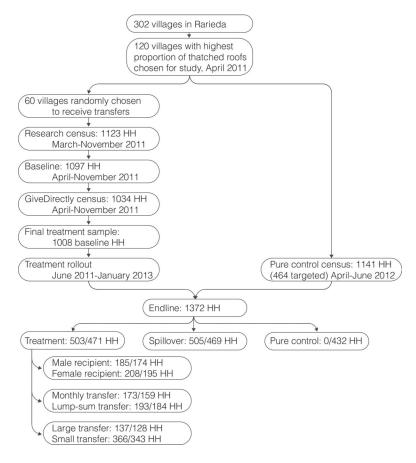
Horowitz Joel L., and Manski Charles F., "Identification and Robustness with Contaminated and Corrupted Data," Econometrica, 63 (1995), 281–302.

Horton Richard, and Smith Richard, "Time to Register for Randomized Trials," British Medical Journal, 319 (1999), 865–866. [PubMed: 10506022]

Jack William, and Suri Tavneet, "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," American Economic Review 104 (2014), 183–223.

Jones Michael P., "Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression," Journal of the American Statistical Association, 91 (1996), 222–230.

Kiecolt-Glaser Janice K., Preacher Kristopher J., MacCallum Robert C., Atkinson Cathie, Malarkey William B., and Glaser Ronald, "Chronic Stress and Age-Related Increases in the Proinflammatory Cytokine IL-6," Proceedings of the National Academy of Sciences, 100 (2003), 9090–9095.

Kirschbaum Clemens, and Hellhammer Dirk H., "Salivary Cortisol in Psychobiological Research: An Overview," Neuropsychobiology, 22 (1989), 150–169. [PubMed: 2485862]

Kirschbaum Clemens, Strasburger Christian J., and Langkrär J, "Attenuated Cortisol Response to Psychological Stress but Not to CRH or Ergometry in Young Habitual Smokers," Pharmacology Biochemistry and Behavior, 44 (1993), 527–531.

Kling Jeffrey R, Liebman Jeffrey B, and Katz Lawrence F, "Experimental Analysis of Neighborhood Effects," Econometrica, 75 (2007), 83–119.

Knorr Ulla, Vinberg Maj, Kessing Lars V., and Weeterslev Jorn, "Salivary Cortisol in Depressed Patients versus Control Persons: A Systematic Review and Meta-Analysis," Psychoneuroendocrinology, 35 (2010), 1275–1286. [PubMed: 20447770]

Kuhn Peter, Kooreman Peter, Soetevent Adriaan, and Kapteyn Arie, "The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery," American Economic Review, 101 (2011), 2226–2247.

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

Lund Crick, Mary De Silva Sophie Plagerson, Cooper Sara, Chisholm Dan, Das Jishnu, Knapp Martin, and Patel Vikram, "Poverty and Mental Disorders: Breaking the Cycle in Low-Income and Middle-Income Countries," Lancet, 378 (2011), 1502–1514. [PubMed: 22008425]

Lundberg Shelly J., Pollak Robert A., and Wales Terence J., "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit," Journal of Human Resources, 32 (1997), 463–480.

MacKinnon James G., and Magee Lonnie, "Transforming the Dependent Variable in Regression Models," International Economic Review, 31(1990), 315–339.

Macours Karen, Schady Norbert, and Vakis Renos, "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment," American Economic Journal: Applied Economics, 4 (2012), 247–373.

Maluccio John, "The Impact of Conditional Cash Transfers on Consumption and Investment in Nicaragua," Journal of Development Studies, 46 (2010), 14–38.

Maluccio John, and Flores Rafael, "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Proteccion Social," Research Report 141, International Food Policy Research Institute (IFPRI), 2005.

Maluccio John, Murphy Alexis, and Regalia Ferdinando, "Does Supply Matter? Initial Schooling Conditions and the Effectiveness of Conditional Cash Transfers for Grade Progression in Nicaragua," Journal of Development Effectiveness, 2 (2010), 87–116.

McKenzie David J., "Beyond Baseline and Follow-Up: The Case for More T in Experiments," Journal of Development Economics, 99 (2012), 210–221.

———, "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition," Policy Research Working Paper 7391, World Bank, 2015.

Mendolia Silvia, "The Impact of Husband's Job Loss on Partner's Mental Health," Review of Economics of the Household, 12 (2014), 277–294.

Modigliani Franco, and Brumberg Richard, "Utility Analysis and the Consumption Function," in Post-Keynesian Economics, Kurihara Kenneth, ed. (New Brunswick, NJ: Rutgers University Press, 1954).

Oswald Andrew J., Proto Eugenio, and Sgroi Daniel, "Happiness and Productivity," University of Warwick Working Paper, 2009.

Ozer Emily, Fernald Lia, Weber Ann, Flynn Emily, and Tyler VanderWeele, "Does Alleviating Poverty Affect Mothers' Depressive Symptoms? A Quasi-Experimental Investigation of Mexico's Oportunidades Programme," International Journal of Epidemiology, 40 (2011), 1565–1576. [PubMed: 21737404]

Paxson Christina, and Schady Norbert, "Does Money Matter? The Effects of Cash Transfers on Child Development in Rural Ecuador," Economic Development and Cultural Change, 59 (2010), 187–229. [PubMed: 20821896]

Pence Karen M., "The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving," Contributions to Economic Analysis and Policy, 5 (2006), 1–26. [PubMed: 21339848]

Pepper John V., "Robust Inferences from Random Clustered Samples: An Application Using Data from the Panel Study of Income Dynamics," Economics Letters, 75 (2002), 341–345.

Posel Dorrit, Fairburn James A., and Lund Frances, "Labour Migration and Households: A Reconsideration of the Effects of the Social Pension on Labour Supply in South Africa," Economic Modelling, 23 (2006), 836–853.

Radloff Lenore S., "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population," Applied Psychological Measurement, 1 (1977), 385–401.

Rosenberg Morris, Society and the Adolescent Self-Image (Princeton, NJ: Princeton University Press, 1965).

Rosenthal Robert, "The File Drawer Problem and Tolerance for Null Results," Psychological Bulletin, 86 (1979), 638–641.

Rosero Jose, and Oosterbeek Hessel, "Trade-Offs between Different Early Childhood Interventions: Evidence from Ecuador," Tinbergen Institute Discussion Paper 11–102/3, Tinbergen Institute, 2011.

Ross Russell, "Atherosclerosis An Inflammatory Disease," New England Journal of Medicine, 340 (1999), 115–126. [PubMed: 9887164]

Rotter Julian B., "Generalized Expectancies for Internal versus External Control of Reinforcement," Psychological Monographs: General and Applied, 80 (1966), 1–28.

Sadoulet Elisabeth, de Janvry Alain, and Davis Benjamin, "Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico," World Development, 29 (2001), 1043–1056.

Schady Norbert, and Rosero José, "Are Cash Transfers Made to Women Spent Like Other Sources of Income?," Economics Letters, 101 (2008), 246–248.

Scheier Michael F., Carver Charles S., and Bridges Michael W., "Distinguishing Optimism from Neuroticism (and Trait Anxiety, Self-Mastery, and Self-Esteem), A Reevaluation of the Life Orientation Test," Journal of Personality and Social Psychology, 67 (1994), 1063. [PubMed: 7815302]

Simes Robert John, "Publication Bias: The Case for and International Registry of Clinical Trials," Journal of Clinical Oncology, 4 (1986), 1529–1541. [PubMed: 3760920]

Skoufias Emmanuel, Unar Mishel, and de Cossio Teresa Gonzalez, "The Poverty Impacts of Cash and In-Kind Transfers: Experimental Evidence from Rural Mexico," Journal of Development Effectiveness, 5 (2013), 401–429.

Ssewamala Fred, Han Chang-Keun, and Neilands Torsten, "Asset Ownership and Health and Mental Health Functioning among AIDS-Orphaned Adolescents: Findings from a Randomized Clinical Trial in Rural Uganda," Social Science & Medicine, 69 (2009), 191–198. [PubMed: 19520472]

Ssewamala Fred, Neilands Torsten, Waldfogel Jane, and Ismayilova Leyla, "The Impact of a Comprehensive Microfinance Intervention on Depression Levels of AIDS-Orphaned Children in Uganda," Journal of Adolescent Health, 50 (2012), 346–352. [PubMed: 22443837]

Steptoe Andrew, Feldman Pamela J., Kunz Sabine, Owen Natalie, Willemsen Gonneke, and Marmot Michael, "Stress Responsivity and Socioeconomic Status: A Mechanism for Increased

Cardiovascular Disease Risk?," European Heart Journal, 23 (2002), 1757–1763. [PubMed: 12419295]

Steptoe Andrew, Willemsen Gonneke, Owen Natalie, Flower Louise, and Mohamed-Ali Vidya, "Acute Mental Stress Elicits Delayed Increases in Circulating Inflammatory Cytokine Levels," Clinical Science, 101 (2001), 185–192. [PubMed: 11473494]

Straub Rainer H., "Bottom-Up and Top-Down Signaling of IL-6 with and without Habituation?," Brain, Behavior, and Immunity, 20 (2006), 37–39.

Thomas Duncan, "Intra-Household Resource Allocation: An Inferential Approach," Journal of Human Resources, 25 (1990), 635–664.

Townsend Richard, "Risk and Insurance in Village India," Econometrica, 62 (1994), 539–591.

Westfall Peter H., Young S. Stanley, and Wright S. Paul, "On Adjusting P-Values for Multiplicity," Biometrics, 49 (1993), 941–945.

Wilckens Thomas, "Glucocorticoids and Immune Function: Physiological Relevance and Pathogenic Potential of Hormonal Dysfunction," Trends in Pharmacological Sciences, 16 (1995), 193–197. [PubMed: 7652928]

Zwane Alix, Zinman Jonathan, Eric Van Dusen William Pariente, Null Clair, Miguel Edward, Kremer Michael, Karlan Dean, Hornbeck Richard, Xavier Giná Esther Duflo, Devoto Florencia, Crepon Bruno, and Banerjee Abhijit, "Being Surveyed Can Change Later Behavior and Related Parameter Estimates," Proceedings of the National Academy of Sciences, 108 (2011), 1821–1826.

FIGURE I.

Timeline of Study

Timeline and treatment arms. Numbers with slashes designate baseline/endline number of households in each treatment arm. Male versus female recipient was randomized only for households with cohabitating couples. Large transfers were administered by making additional transfers to households that had previously been assigned to treatment. The lump-sum versus monthly comparison is restricted to small transfer recipient households.

TABLE I

Baseline Differences in Index Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Control mean (std. dev.)	Treatment effect	Female recipient	Monthly transfer	Large transfer	N
Value of nonland assets (US$)	383.36 (374.15)	−1.15 (24.74) [1.00]	15.53 (43.62) [0.92]	25.16 (39.33) [0.91]	13.76 (42.77) [0.99]	1,008
Nondurable expenditure (US$)	181.99 (127.16)	−6.16 (8.31) [0.98]	−28.05* (15.14) [0.31]	−8.01 (13.28) [0.91]	−5.56 (14.36) [0.99]	1,008
Total revenue, monthly (US$)	84.92 (402.59)	−33.19* (18.54) [0.43]	−31.77** (14.34) [0.15]	−7.59 (14.99) [0.91]	−10.77 (12.38) [0.95]	1,008
Food security index	0.00 (1.00)	0.00 (0.06) [1.00]	0.05 (0.09) [0.92]	0.25** (0.10) [0.08]*	−0.01 (0.09) [0.99]	1,008
Health index	0.01 (1.02)	0.03 (0.06) [0.98]	0.26*** (0.09) [0.04]**	0.14 (0.10) [0.54]	−0.14 (0.10) [0.69]	1,008
Education index	0.00 (1.00)	−0.07 (0.06) [0.89]	0.14 (0.09) [0.41]	0.16* (0.09) [0.42]	−0.05 (0.09) [0.98]	853
Psychological well-being index	0.00 (1.00)	0.03 (0.05) [0.98]	0.02 (0.08) [0.92]	0.19** (0.08) [0.13]	0.18** (0.08) [0.17]	1,569
Female empowerment index	0.00 (1.00)	−0.05 (0.07) [0.98]	0.08 (0.11) [0.92]	0.18 (0.12) [0.52]	0.03 (0.12) [0.99]	751
Joint test (*p*-value)		.64	.00***	.02**	.36	

Notes. OLS estimates of baseline differences in treatment arms. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. FWER-corrected *p*-values are shown in brackets. Column (1) reports the mean and standard deviation of the control group for a given outcome variable. Column (2) reports the basic treatment effect, that is, comparing treatment households to control households within villages. Column (3) reports the relative treatment effect of transferring to the female compared to the male; column (4) the relative effect of monthly compared to lump-sum transfers; and column (5) that of large compared to small transfers. The unit of observation is the household for all outcome variables except for the psychological variables index, where it is the individual. The sample is restricted to cohabitating couples for the female empowerment index and households with school-age children for the education index. The comparison of monthly to lump-sum transfers excludes large transfer recipient households, and that for male versus female recipients excludes single-headed households. All columns include village-level fixed effects and cluster standard errors at the household level. The last row shows joint significance of the coefficients in the corresponding column from SUR estimation.

*
denotes significance at 10%

**
5%, and

***
1% levels.

TABLE II

Treatment Effects: Index Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Control mean (std. dev.)	**Treatment effect**	**Female recipient**	**Monthly transfer**	**Large transfer**	**N**
Value of nonland assets (US$)	494.80 (415.32)	301.51 *** (27.25) [0.00] ***	−79.46 (50.38) [0.52]	−91.85 ** (45.92) [0.28]	279.18 *** (49.09) [0.00] ***	940
Nondurable expenditure (US$)	157.61 (82.18)	35.66 *** (5.85) [0.00] ***	−2.00 (10.28) [0.92]	−4.20 (10.71) [0.99]	21.25 ** (10.49) [0.22]	940
Total revenue, monthly (US$)	48.98 (90.52)	16.15 *** (5.88) [0.02] **	5.41 (10.61) [0.92]	16.33 (11.07) [0.59]	−2.44 (8.87) [0.84]	940
Food security index	0.00 (1.00)	0.26 *** (0.06) [0.00] ***	0.06 (0.09) [0.92]	0.26 ** (0.11) [0.13]	0.18 * (0.10) [0.25]	940
Health index	0.00 (1.00)	−0.03 (0.06) [0.82]	0.10 (0.09) [0.72]	0.01 (0.10) [0.99]	−0.09 (0.09) [0.72]	940
Education index	0.00 (1.00)	0.08 (0.06) [0.43]	0.06 (0.09) [0.92]	−0.05 (0.10) [0.99]	0.05 (0.09) [0.84]	823
Psychological well-being index	0.00 (1.00)	0.26 *** (0.05) [0.00] ***	0.14 * (0.08) [0.43]	0.01 (0.08) [0.99]	0.26 *** (0.08) [0.00] ***	1,474
Female empowerment index	0.00 (1.00)	−0.01 (0.07) [0.88]	0.17 * (0.10) [0.51]	0.05 (0.12) [0.99]	0.22 ** (0.11) [0.22]	698
Joint test (p-value)		.00 ***	.11	.04 **	.00 ***	

Notes. OLS estimates of treatment effects. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. Higher values correspond to "positive" outcomes. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. FWER-corrected *p*-values are shown in brackets. Column (1) reports the mean and standard deviation of the spillover group, column (2) the basic treatment effect, that is, comparing treatment households to control households within villages. Column (3) reports the relative treatment effect of transferring to the female compared to the male; column (4) the relative effect of monthly compared to lump-sum transfers; and column (5) that of large compared to small transfers. The unit of observation is the household for all outcome variables except for the psychological variables index, where it is the individual. The sample is restricted to cohabitating couples for the female empowerment index, and households with school-age children for the education index. The comparison of monthly to lump-sum transfers excludes large transfer recipient households, and that for male versus female recipients excludes single-headed households. All columns include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level. The last row shows joint significance of the coefficients in the corresponding column from SUR estimation.

*
denotes significance at 10%

**
5%, and

***
1% levels.

TABLE III

Spillover Effects: Index Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
			Spillover effects				Lee bounds		Horowitz-Manski bounds	
	All HH estimate	All HH estimate	Thatched estimate	Thatched estimate	Test (1) = (3) *p*-value	Test (2) = (4) *p*-value	Lower	Upper	Lower	Upper
Includes controls	No	Yes	No	Yes	No	Yes	No	No	No	No
Value of nonland assets (US$)	1.00 (21.44)	−11.99 (19.98)	−18.73 (21.14)	−32.61 (19.76)	.01 ***	.00 ***	−3.38 (17.62)	12.84 (18.68)	−2.38 (19.92)	7.39 (20.06)
Nondurable expenditure (US$)	−7.77 (7.20)	−11.89 * (6.50)	−7.31 (7.27)	−12.21 * (6.67)	.82	.88	−9.47 (6.36)	−4.08 (6.37)	−8.93 (5.79)	−5.86 (5.83)
Total revenue, monthly (US$)	−3.68 (6.18)	−3.64 (6.35)	−5.23 (5.84)	−5.78 (6.01)	.56	.45	−4.29 (5.67)	2.32 (5.60)	−4.18 (6.18)	−1.91 (6.22)
Food security index	0.06 (0.09)	0.05 (0.09)	0.06 (0.10)	0.05 (0.10)	.93	.90	−0.01 (0.08)	0.08 (0.08)	0.03 (0.08)	0.07 (0.08)
Health index	−0.06 (0.08)	−0.07 (0.08)	−0.06 (0.08)	−0.08 (0.08)	.80	.66	−0.10 (0.07)	−0.03 (0.07)	−0.07 (0.07)	−0.04 (0.07)
Education index	0.01 (0.07)	−0.01 (0.06)	−0.00 (0.08)	−0.03 (0.07)	.36	.29	−0.16 ** (0.08)	0.06 (0.08)	−0.01 (0.07)	0.03 (0.07)
Psychological well-being index	0.11 (0.07)	0.10 (0.07)	0.11 (0.07)	0.10 (0.07)	.92	.95	−0.01 (0.05)	0.07 (0.05)	0.01 (0.05)	0.05 (0.05)
Female empowerment index	0.21 ** (0.09)	0.21 ** (0.08)	0.23 ** (0.09)	0.22 ** (0.09)	.50	.50	0.15 * (0.08)	0.25 *** (0.08)	0.18 ** (0.08)	0.23 *** (0.08)
Joint test (*p*-value)	.38	.25	.23	.11						

Notes. OLS estimates of spillover effects. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. The unit of observation is the household for all variables except psychological well-being, where it is the individual. The sample includes all households and individuals, except for the intrahousehold index, where it is restricted to cohabitating couples, and for the education index, where it is restricted to households with school-age children. Columns (1) and (2) report the "naïve" estimate of spillover effects, including spillover households that upgraded to metal roofs between baseline and endline. Columns (3) and (4) report estimates of the spillover effect excluding metal-roof households. Columns (1) and (3) exclude baseline covariates. Columns (2) and (4) include baseline covariates. Column (5) reports the *p*-value of the equality for the coefficient estimates in (1) and (3) after joint estimation of the two models using SUR. Column (6) reports the *p*-value of the equality for the coefficient estimates in (2) and (4) after joint estimation of the two models using SUR. The last row reports *p*-values on the joint significance of all coefficients in a given column after joint estimation using SUR. Columns (7) and (8) report the lower and upper Lee bounds adjusting for differential "attrition" of five spillover households into metal-roof upgrade. Columns (9) and (10) report lower and upper Horowitz-Manski bounds, imputing outcomes for the five attriting households using the 95th and 5th percentiles of observed outcomes, respectively. The sample is restricted as in the previous tables. In columns (1)–(4), standard errors clustered at the village level are reported in parentheses. In columns (7)–(10), bootstrapped standard errors are reported in parentheses.

*
denotes significance at 10%

**
5%, and

***
1% levels.

TABLE IV

Treatment Effects: Psychological Well-Being

	(1)	(2)	(3)	(4)	(5)	(6)
	Control mean (std. dev.)	Treatment effect	Female recipient	Monthly transfer	Large transfer	N
Log cortisol (no controls)	2.46 (0.89)	0.00 (0.05)	−0.17** (0.07)	0.16* (0.08)	−0.09 (0.07)	1,456
Log cortisol (with controls)	−0.04 (0.88)	0.01 (0.05)	−0.17** (0.07)	0.17** (0.08)	−0.12* (0.07)	1,456
Depression (CESD)	26.48 (9.31)	−1.16*** (0.44)	−0.77 (0.67)	−1.40* (0.73)	−1.22* (0.68)	1,474
Worries	0.00 (1.00)	−0.13*** (0.05)	−0.04 (0.07)	−0.11 (0.08)	−0.07 (0.08)	1,474
Stress (Cohen)	0.00 (1.00)	−0.26*** (0.05)	−0.02 (0.08)	−0.02 (0.09)	−0.24*** (0.08)	1,474
Happiness (WVS)	0.00 (1.00)	0.16*** (0.05)	0.07 (0.08)	0.03 (0.09)	0.07 (0.08)	1,474
Life satisfaction (WVS)	0.00 (1.00)	0.17*** (0.05)	−0.07 (0.07)	0.12 (0.08)	0.19** (0.08)	1,474
Trust (WVS)	0.00 (1.00)	0.04 (0.05)	0.08 (0.08)	−0.08 (0.08)	−0.04 (0.08)	1,474
Locus of control	0.00 (1.00)	0.03 (0.05)	0.04 (0.08)	−0.03 (0.09)	0.08 (0.08)	1,474
Optimism (Scheier)	0.00 (1.00)	0.10* (0.05)	0.07 (0.08)	0.02 (0.09)	0.16* (0.09)	1,474
Self-esteem (Rosenberg)	0.00 (1.00)	0.00 (0.05)	0.19** (0.09)	0.09 (0.09)	−0.15 (0.10)	1,474
Psychological well-being index	0.00 (1.00)	0.26*** (0.05)	0.14* (0.08)	0.01 (0.08)	0.26*** (0.08)	1,474
Joint test (*p*-value)		.00***	.21	.21	.00***	

Notes. OLS estimates of treatment and spillover effects. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. All variables are coded in *z*-score units, except raw cortisol, which is coded in nmol/L, and depression, which is coded in points. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. Column (1) reports the mean and standard deviation of the control group for a given outcome variable. column (2) reports the basic treatment effect, that is, comparing treatment households to control households within villages. Column (3) reports the relative treatment effect of transferring to the female compared to the male; column (4) the relative effect of monthly compared to lump-sum transfers; and column (5) that of large compared to small transfers. The unit of observation is the individual. The comparison of monthly to lump-sum transfers excludes large transfer recipient households, and that for male versus female recipients excludes single-headed households. All columns include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level. The last row shows joint significance of the coefficients in the corresponding column from SUR estimation.

*
denotes significance at 10%

**
5%, and

***
1% level.

TABLE V

Treatment Effects: Consumption

	(1)	(2)	(3)	(4)	(5)	(6)
	Control mean (std. dev.)	**Treatment effect**	**Female recipient**	**Monthly transfer**	**Large transfer**	**N**
Food total (US$)	104.46 (58.50)	19.46***(4.19)	−1.81 (7.37)	1.79 (7.42)	8.28 (7.59)	940
Cereals (US$)	22.55 (17.18)	2.23**(1.13)	0.37 (1.87)	−1.06 (1.86)	2.68 (2.07)	940
Meat & fish (US$)	12.97 (13.75)	5.05***(1.01)	0.87 (1.82)	−2.93 (1.92)	2.52 (1.63)	940
Alcohol (US$)	6.38 (16.56)	−0.93 (0.99)	1.56 (1.62)	1.03 (1.64)	−1.42 (1.33)	940
Tobacco (US$)	1.52 (4.13)	−0.15 (0.22)	0.12 (0.34)	0.42 (0.33)	−0.29 (0.30)	940
Social expenditure (US$)	4.36 (5.38)	2.43***(0.48)	−2.06**(0.97)	−0.52 (0.99)	0.62 (0.90)	940
Medical expenditure past month (US$)	6.78 (13.53)	2.58***(0.99)	2.06 (1.86)	−1.34 (1.86)	−0.29 (1.74)	940
Education expenditure (US$)	4.71 (8.68)	1.08**(0.51)	0.48 (0.88)	−0.02 (0.87)	1.15 (0.91)	940
Non-durable expenditure (US$)	157.61 (82.18)	35.66***(5.85)	−2.00 (10.28)	−4.20 (10.71)	21.25**(10.49)	940
Joint test (*p*-value)		.00***	.47	.13	.01***	

Notes. OLS estimates of treatment and spillover effects. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. All variables are reported in PPP-adjusted US$ and are top-coded for the highest 1% of observations. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. Column (1) reports the mean and standard deviation of the control group for a given outcome variable. Column (2) reports the basic treatment effect, that is, comparing treatment households to control households within villages. Column (3) reports the relative treatment effect of transferring to the female compared to the male; column (4) the relative effect of monthly compared to lump-sum transfers; and column (5) that of large compared to small transfers. The unit of observation is the household. The comparison of monthly to lump-sum transfers excludes large transfer recipient households, and that for male versus female recipients excludes single-headed households. All columns include village-level fixed effects and control for baseline outcomes. The last row shows joint significance of the coefficients in the corresponding column from SUR estimation.

*
denotes significance at 10%

**
5%, and

***
1% levels.

TABLE VI

Treatment Effects: Assets and Agricultural and Business Activities

	(1) Control mean (std. dev.)	(2) Treatment effect	(3) Female recipient	(4) Monthly transfer	(5) Large transfer	(6) N
Panel A: assets						
Value of nonland assets (US$)	494.80 (415.32)	301.51*** (27.25)	−79.46 (50.38)	−91.85** (45.92)	279.18*** (49.09)	940
Value of livestock (US$)	166.82 (240.59)	83.18*** (15.22)	4.84 (29.32)	0.08 (27.36)	63.45** (28.51)	940
Value of durable goods (US$)	207.30 (130.60)	52.59*** (8.61)	−0.24 (14.40)	−7.31 (14.16)	64.90*** (15.70)	940
Value of savings (US$)	10.93 (29.09)	10.10*** (2.46)	−3.31 (5.03)	1.86 (4.57)	10.26** (5.04)	940
Land owned (acres)	1.31 (1.88)	0.04 (0.14)	−0.08 (0.18)	0.04 (0.17)	0.35 (0.32)	940
Has nonthatched roof (dummy)	0.16 (0.37)	0.24*** (0.03)	−0.11** (0.05)	−0.12** (0.05)	0.23*** (0.05)	940
Joint test (*p*-value)		.00***	.29	.22	.00***	
Panel B: business activities						
Wage labor primary income (dummy)	0.16 (0.37)	0.00 (0.02)	0.02 (0.04)	0.02 (0.04)	0.01 (0.04)	940
Own farm primary income (dummy)	0.56 (0.50)	−0.01 (0.03)	−0.00 (0.05)	−0.00 (0.05)	0.01 (0.05)	940
Nonagricultural business owner (dummy)	0.32 (0.47)	0.01 (0.03)	−0.02 (0.05)	0.08 (0.05)	0.01 (0.05)	940
Total revenue, monthly (US$)	48.98 (90.52)	16.15*** (5.88)	5.41 (10.61)	16.33 (11.07)	−2.44 (8.87)	940
Total expenses, monthly (US$)	23.95 (61.71)	12.53*** (4.21)	5.42 (7.45)	9.41 (7.75)	−0.35 (6.23)	940
Total profit, monthly (US$)	20.78 (46.22)	−0.21 (3.68)	1.41 (6.68)	7.29 (7.92)	−2.02 (5.32)	940
Joint test (*p*-value)		0.00***	0.90	0.59	1.00	

Notes. OLS estimates of treatment and spillover effects. Outcome variables are listed on the left, and described in detail in Appendix Table A.2. Variables are in PPP-adjusted 2012 $ and are top-coded for the highest 1% of observations. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. Column (1) reports the mean and standard deviation of the control group for a given outcome variable. Column (2) reports the basic treatment effect, that is, comparing treatment households to control households within villages. Column (3) reports the relative treatment effect of transferring to the female compared to the male; column (4) the relative effect of monthly compared to lump-sum transfers; and column (5) that of large compared to small transfers. The unit of observation is the household. The comparison of monthly to lump-sum transfers excludes large transfer recipient households, and that for male versus female recipients excludes single-headed households. All columns except the spillover regressions include village-level fixed effects and control for baseline outcomes. The last row shows joint significance of the coefficients in the corresponding column from SUR estimation.

*
denotes significance at 10%

**
5%, and

***
1% levels.