Household Response to Income Changes: Evidence from an Unconditional Cash Transfer Program in Kenya*

Johannes Haushofer†, Jeremy Shapiro‡

January 27, 2014

Abstract

This paper studies the response of poor rural households in rural Kenya to large temporary income changes. Using a randomized controlled trial, households were randomly assigned to receive unconditional cash transfers of at least USD 404 from the NGO *GiveDirectly*. We designed the experiment to address several long-standing questions in the economics literature: what is the shape of households' Engel curves? Do household members effectively pool income? Are there constraints to savings? Do transfers generate externalities? In addition, we study in detail the effects of transfers on psychological well-being and levels of the stress hormone cortisol. We randomized at both the village and household levels; further, within the treatment group, we randomized recipient gender (wife vs. husband), transfer timing (lump-sum transfer vs. monthly installments over 9 months), and transfer magnitude (USD 404 vs. USD 1,520). We find a strong consumption response to transfers, with an increase in monthly consumption from USD 157 to USD 194 four months after the transfer ended. Implied expenditure elasticities for food, medical and education expenditures range between 0.84 and 1.47, while the point estimates are negative for alcohol and tobacco. Intriguingly, recipient gender does not affect the household response to the program. Households may face savings constraints: monthly transfers are more likely than lump-sum transfers to improve food security, while lump-sum transfers are more likely to be spent on durables. We find no evidence for externalities on non-recipients except for a significant positive spillover on female empowerment. Transfer recipients experience large increases in psychological well-being, and several types of transfers lead to reductions in levels of the stress hormone cortisol. Together, these results suggest that unconditional cash transfers have significant impacts on consumption and psychological well-being.

*We are deeply grateful to Faizan Diwan for outstanding project management. We further thank the study participants for generously giving their time; Marie Collins, Conor Hughes, Chaning Jang, Bena Mwongeli, Joseph Njoroge, Kenneth Okumu, James Vancel, and Matthew White for excellent research assistance; the team of GiveDirectly (Michael Faye, Raphael Gitau, Piali Mukhopadhyay, Paul Niehaus, Joy Sun, Carolina Toth, Rohit Wanchoo) for fruitful collaboration; Petra Persson for designing the intrahousehold bargaining and domestic violence module; and Anna Aizer, Michael Anderson, Abhijit Banerjee, Victoria Baranov, Dan Björkegren, Chris Blattman, Kate Casey, Arun Chandrasekhar, Michael Clemens, Rebecca Dizon-Ross, Esther Duflo, Simon Galle, Rachel Glennerster, Ben Golub, Nina Harari, Anil Jain, Anna Folke Larsen, Helene Bie Lilleør, David McKenzie, Paul Niehaus, Dina Pomeranz, Vincent Pons, Tristan Reed, Nick Ryan, Emma Rothschild, Simone Schaner, Xiao-Yu Wang, and seminar participants at MIT, Harvard, and NEUDC for comments and discussion. All errors are our own. This research was supported by NIH Grant R01AG039297 and Cogito Foundation Grant R-116/10 to Johannes Haushofer.

†Abdul Latif Jameel Poverty Action Lab, MIT, E53-389, 30 Wadsworth St., Cambridge, MA 02142. joha@mit.edu

‡PhD (MIT). Shapiro is a co-founder and former director of *GiveDirectly, Inc.* (2009–2012). This paper does not necessarily represent the views of *GiveDirectly, Inc.* jeremypshapiro@gmail.com

1 Introduction

The response of poor households to income changes is of critical interest both to academics and policy makers. It is a crucial element in modeling the consumption and savings choices of households, and a central ingredient in designing tax and transfers policy, labor market policy, and insurance markets (Deaton 1992; Hall and Mishkin 1982; Jappelli and Pistaferri 2010). In developing countries, it can inform the design of consumption support policies and redistribution programs. In studying the household response to income changes, academics and policy makers are increasingly interested not only in the effects on consumption, labor supply, and wealth, but also more general measures of welfare such as psychological well-being.

However, estimating such responses from observational data alone presents significant challenges. In the cross-section, households that have different resources have different tastes, different opportunities, and probably face different prices, which complicates the interpretation of cross-sectional estimates of elasticities. In the time series, changes in income are typically accompanied with changes in the economic environment faced by the household (e.g. changes in wages or the productivity of labor). Finally, because policy makers in developing countries tend to be wary of unconditional income transfers, most income redistribution to the poor is either in kind or attached to conditionalities, and therefore few natural experiments exist. Indeed, Heckman (1992) praised the early social experiments (such as the Negative Income Tax Experiments in the US) for distinguishing income and substitution effects from higher wages, precisely because this is one of the few cases where it is difficult to think of a substitute for an experiment.

Here we study the response of households to income changes using data from a randomized controlled trial (RCT) of a large, one-time, unanticipated unconditional cash transfer. Between 2011 and 2013, the NGO *GiveDirectly* sent unconditional cash transfers of at least USD 404[1], or at least twice the monthly average household consumption in the area, to randomly chosen poor households in Kenya through the mobile money system *M-Pesa*. The transfers were explicitly described to households as fully unconditional, and as short-term windfalls (in one lump sump or monthly installment over 9 months), rather than as a promise of recurring payments for the long term. We surveyed randomly selected treatment and control households both before the program and between 1 and 14 months after it ended.

The experiment was designed to assess not only the overall impact of such transfers, but also to answer several longstanding questions in the economics literature: what is the shape of households' Engel curves? What is households' effective discount rate? Do household members effectively pool income? Do they face constraints to savings? Do transfers generate significant externalities (positive or negative) on non-beneficiaries? And finally, do transfers affect psychological well-being and levels of the stress hormone cortisol? To answer these questions, we carried out a two stage randomization, at the village and household level. Further, within the treatment group, we randomized the transfer

[1] All USD values are calculated at purchasing power parity, using is the 2012 World Bank PPP estimate for private consumption in Kenya: 0.016.

recipient within the household (wife vs. husband), the transfer timing (monthly installments over nine months vs. one-time lump sum transfer) and transfer magnitude (USD 404 vs. USD 1,520). To study the response of consumption over time, we also randomized survey timing from 1 to 14 months after the end of transfers disbursements.

To complement this ambitious experimental design, we collected detailed data on a broad range of outcomes for 1440 households (1372 at endline). The survey instruments included modules for consumption, asset holdings, self-employment activities and earnings, health, education, food security, female empowerment, and psychological well-being. A novel feature was the collection of a biomarker of stress, i.e. levels of the stress hormone cortisol. In addition, we administered a village-level questionnaire to capture general equilibrium effects. Because of the large number of outcomes, we address issues of multiple inference by pre-specifying the basic reduced form analysis (the pre-analysis plan is available at www.socialscienceregistry.org), by grouping the main outcomes into a limited number of index variables for a restricted number of outcome groups, and by adjusting p-values for multiple inference using the family-wise error rate (FWER).

Assessing the increase in permanent income caused by the program would require a full understanding of households' savings, borrowing and investment behavior and opportunities, and knowledge of what they expected regarding future transfers, and is beyond the scope of this paper. However, when households were surveyed, on average 4.3 months after the end of the program, we observe an increase in monthly non-durable expenditure of USD 36 relative to a control group mean of USD 157. Under standard time-separability assumptions, we can use the transfer as an instrument for the change in non-durable expenditure and thus estimate the elasticity of different expenditure heads with respect to total expenditure. We find elasticities of food expenditures of 0.83, medical expenditure of 1.47, education expenditure of 0.84, and social expenditure of 1.60. The elasticities for alcohol and tobacco expenditure are negative and insignificant. The elasticities implied by large and small transfers are similar overall, suggesting relatively constant elasticities. While equality of the IV and OLS estimates can be rejected in some cases, the two are generally similar. Food security increased significantly for transfer recipients, and we observe an increase in female empowerment at the village level. In contrast, education and health outcomes were not affected by transfers.

Households invested part of the transfers in durables and assets for their self-employment activities. We find a significant increase of USD 279 in asset holdings, relative to a control group mean of USD 478. In particular, we find increases in holdings of home durables (notably metal roofs, ownership of which increased by 23 percentage points over a control group mean of 16 percent), and productive assets such as livestock, whose value increases by USD 85 over a control group mean of USD 167. These investments translate into higher revenues from agriculture, animal husbandry, and non-agricultural enterprises; monthly revenue from these sources increases by USD 17 relative to a control group mean of USD 49. Note, however, that this revenue increase is partially offset by an increase in flow expenses for agriculture, animal husbandry, and business (USD 13 relative to a control group mean of USD 24).

Despite the fact that they invest in assets and durables, households appear to be relatively impatient. The impact of the transfer on nondurable consumption declines over time – from USD 29 for households observed in the first three months after the transfer to USD 19 observed on average seven months after the end of the program, which implies an effective annual discount rate of 56 percent. An explanation which helps to reconcile this effect with the investment in assets is that households may find it difficult to save. The comparison between households that received monthly vs. lump-sum transfers is consistent with this interpretation: if households are both credit- and savings-constrained, 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. Conversely, recipients of a lump sum may be keen to invest it immediately into a large durable if they are not sure they can pace their non-durable consumption and save. We find that indeed monthly transfer recipients are significantly less likely to invest in durables such as metal roofs than lump-sum transfer recipients, suggesting that households may be both credit- and savings-constrained. 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 – and even add other money – on their *M-Pesa* account), suggests that the savings constraint at work is more social or behavioral than purely due to lack of access to a savings technology.

Interestingly, we find few differences between female vs. male recipient households in consumption, production, and investment decisions. This result is surprising in light of a large literature suggesting that households may not be unitary, and may thus not pool income (Thomas 1989; Duflo and Udry 2004), although it is consistent with another recent experiment randomizing the gender of the recipient of a non-conditional cash transfer (Benhassine et al. 2013). One possible explanation is that the program did not affect bargaining power because it was explicitly presented as temporary: even so, it suggests a surprising level of insurance between household members, which suggests that these Kenyan households are more efficient than found in Udry (1996) or Duflo and Udry (2004).

A further core contribution of this paper is that it is the first to measure, on a large scale, a biological marker of stress, cortisol, combined with several experimental treatments and several survey measurements of psychological well-being. Overall, we find large and highly significant increases in psychological well-being among transfer recipients; in particular, we document a 0.19 SD increase in happiness, a 0.15 SD increase in life satisfaction, a 0.14 SD reduction in stress, and a significant reduction in depression (all measured by psychological questionnaires). In contrast to most of the traditional economic outcome variables, these indicators showed a treatment effect across the entire distribution of psychological well-being, i.e. both respondents with low and high psychological well-being experienced similar improvements. In addition, cortisol levels show effects across all treatment arms: in particular, we find that cortisol levels 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. These results are particularly intriguing because some of them occur in the absence of effects on other outcomes, both economic and psychological: we find no differences in consumption or savings and investment decisions when transfers are made to the wife vs. the husband, yet we observe significantly lower cortisol levels in both male and female respondents when transfers are made to the wife. One potential explanation for this finding is that unconditional cash transfers are likely to produce heterogeneous treatment effects, which may be best captured by broad outcome measures such as cortisol and psychological well-being, rather than more specific measures of individual dimensions of welfare. Similarly, we observe only few differences in economic outcomes for monthly vs. lump-sum transfers, but significantly lower cortisol levels for lump-sum transfers. One possible explanation for this finding is that monthly transfer recipient households may find it difficult to save their transfers, resulting in stress and increased levels of cortisol. This hypothesis is supported by the consumption decline over time described above. Finally, and more intuitively, we find that large transfers decreased cortisol levels relative to small transfers. Together, these results suggest that cortisol and measures of psychological well-being are useful complements to, and may in fact sometimes be more sensitive than, traditional measures of economic welfare.

This paper contributes to three literatures in economics, public policy, and psychology. First, in economics, this study enables us to rigorously identify the response of households to income changes using a randomized experiment. Previous approaches to this question have used either cross-sectional estimates (Jappelli and Pistaferri 2010; Krueger and Perri 2010; Hall and Mishkin 1982) or time-series data (Deaton and Subramanian 1996; Dynarski et al. 1997; Krueger and Perri 2005; Krueger and Perri 2006; Browning and Crossley 2001) with the attendant concerns about endogeneity and structural assumptions. Another set of studies has used natural or policy shocks to study household behavior (Jensen and Miller 2008a; Dynarski et al. 1997; Kochar 1995; Deaton and Tarozzi 2005; Bodkin 1959; Rosenzweig and Wolpin 1982; Paxson 1993), but such shocks may operate through prices (e.g. in the case of subsidies), may be anticipated by households (in the case of policy changes), may be partly insured, may come bundled with other changes (e.g. insurance), and their effects may not generalize the effects of other types of shocks. The program we study here provides the ideal "helicopter drops of cash" experiment. We find relatively high expenditure elasticities, and to a large extent the experimentally estimated elasticities are similar to those that would have been obtained from cross-sectional estiamtion.

The sub-treatments that were included in the design additionally allow us to address several additional important questions in the development economics literature: randomizing recipient gender allows us to show that households in this sample are surprisingly unitary, in contrast to a number of previous findings suggesting that households may not pool incomes (Udry 1996; Duflo and Udry 2004). In addition, by using both large and small transfer amounts, we can show that the Engel curves for most expenditure categories are approximately linear (Deaton 1992). Finally, by randomizing the timing of transfers (monthly vs. lump-sum), we show that households in this sample are both savings and credit-constrained (Dupas and Robinson 2013a; Ashraf, Karlan, and Yin 2006; Banerjee and Duflo 2005) .

Second, in public policy, this study contributes to the growing empirical literature on the welfare effects of transfers to the poor. Two features of the program are notable in regard. First, the cash transfers we study were targeted at a general poor population sample, chosen simply for meeting a basic means test criterion. In contrast, previous programs focus on particular recipient groups such as micro-entrepreneurs (Blattman, Fiala, and Martinez 2013; De Mel, McKenzie, and Woodruff 2008; Fafchamps et al. 2011), orphans and vulnerable children (Team 2012a; Team 2012b), or pensioners (Duflo 2003). Our results suggest that welfare improvements can be achieved even when transfers are not targeted at such recipient groups. Second, the transfers we study are completely unconditional. Previous studies have shown that asset transfers combined with capacity building and stipends (Banerjee and Duflo 2011; Bandiera et al. 2013), conditional cash transfers (Banerjee et al. 2010; Brune et al. 2011; Bandiera et al. 2013), and unconditional cash transfers (Blattman, Fiala, and Martinez 2013; Cunha, De Giorgi, and Jayachandran 2011) have positive effects on consumption, income, and other welfare measures. However, these programs were rarely entirely unconditional; even in nominally unconditional programs such as Uganda's *Youth Opportunities Program* (Blattman, Fiala, and Martinez 2013), recipients were required to write business plans to receive the transfer, thus creating a clear expectation that the money would be spent on businesses (Devoto et al. 2011; Cunha, De Giorgi, and Jayachandran 2011). In contrast, the *GiveDirectly* cash transfers we study here are completely unconditional; recipients are explicitly told that they are free to spend the transfers however they wish. In this context, our study also contributes to the literature on returns to capital in developing countries. Previous studies have found high rates of return to capital for existing businesses, e.g. and 60 percent for micro-entrepreneurs in Sri Lanka (De Mel, McKenzie, and Woodruff 2008) and 113 percent for shop owners in Kenya (Duflo, Kremer, and Robinson 2008), although estimates of microcredit suggests that the numbers may be lower (Karlan and Zinman 2011). We find substantial increases in consumption for undirected cash transfers, suggesting that the selection of specific population groups as transfer recipients may not always be necessary for transfers to be effective.

Finally, we contribute to an emerging literature on the psychology and neurobiology of poverty. Recent work has suggested that poverty may have particular psychological and neurobiological consequences, and that these, in turn, may affect economic choice in a potentially disadvantageous fashion (Mani et al. 2013; Chemin et al. 2013; Shah, Mullainathan, and Shafir 2012; Haushofer, Fehr, and Schunk 2013; Haushofer et al. 2011; Haushofer and Fehr 2013). A particular version of this hypothesis suggests that poverty may cause stress, and stress may affect economic behavior by increasing discount rates (Haushofer et al. 2011; Haushofer 2011; Haushofer, Fehr, and Schunk 2013; Cornelisse et al. 2013). In previous work, we have found that negative income shocks lead to increases in levels of the stress hormone cortisol among Kenyan famers (Chemin et al. 2013), and that pharmacological administration of the cortisol precursor hydrocortisone increases discount rates (Cornelisse et al. 2013). We have further shown that similar behavioral effects result when individuals suffer large negative income shocks (Haushofer, Fehr, and Schunk 2013). The present study fills an important gap in this proposed feedback loop: despite a number of studies showing correlations between poverty and psychological outcomes (Stevenson and Wolfers 2008; Sacks, Stevenson, and Wolfers 2012; Kahneman and Deaton 2010; Haushofer and Fehr 2013; Haushofer, Fehr, and Schunk 2013) as well as poverty and cortisol levels (Haushofer et al. 2011; Cohen et al. 2006; Cohen, Doyle, and Baum 2006), causal evidence on these relationships is scarce (Arnetz et al. 1991; Fernald and Gunnar 2009). To our knowledge, the present study is the first to rigorously identify the effect of a decrease in poverty on cortisol levels, and one of the first to study the effects of poverty alleviation on psychological well-being (Baird, De Hoop, and Özler 2013; Kling, Liebman, and Katz 2007; Devoto et al. 2011)[2]. We find large increases in psychological well-being and reductions in stress as a result of transfers, lending support to the proposed relationship between poverty and psychological outcomes. In addition, we find significant reductions in cortisol levels in several treatment arms: specifically, large transfers, transfers to women, and lump-sum transfers lead to significantly lower cortisol levels than small transfers, transfers to men, and monthly transfers. Some of these effects occur in the absence of differences in traditional outcome variables. Together, these results support a causal effect of poverty (alleviation) on (reductions in) stress levels. More broadly, they suggest that psychological well-being and cortisol can complement traditional welfare measures, and in some cases may in fact respond to interventions with greater sensitivity than these traditional measures.

The remainder of the paper is organized as follows. Sections 2 and 3 describe the *GiveDirectly* program and the evaluation design. Section 4 summarizes the reduced form impacts of the program on all outcomes, including psychological well-being and cortisol levels. Section 5 presents detailed results on the consumption response to the program; we first estimate elasiticies for a variety of consumption goods in Section 5.1, and then estimate returns to investments and households' discount factor in Section 5.2. Section 6 concludes.

2 The *GiveDirectly* Unconditional Cash Transfer Program

GiveDirectly, Inc. (GD; www.givedirectly.org) is an international NGO founded in 2010, whose mission is to make unconditional cash transfers to poor households in developing countries. We note that Jeremy Shapiro, an author of this study, is a co-founder and former Director of *GiveDirectly* (2009–2012). It began operations in Kenya in 2011 (Goldstein 2013). *GD* selects poor households by first identifying poor regions of Kenya according to census data. In the case of the present study, the region chosen was Rarieda, a peninsula in Lake Victoria west of Kisumu in Western Kenya. Following the choice of a region in which to operate, *GD* identifies target villages. In the case of Rarieda, this was achieved through a rough estimation of the population of villages and the proportion of households lacking a metal roof, which is *GD*'s targeting criterion. The criterion was established by *GD* in prior work as an objective and highly predictive indicator of poverty.

[2] Note that the second part of the proposed loop, i.e. that from stress to decision-making, cannot be aptly tested with an economic intervention like unconditional cash transfers, because such interventions may change economic choice through channels other than stress (e.g. the budget constraint). The link between stress and economic choice is therefore best studied in laboratory settings where stress can be manipulated independently of economic variables.

Villages with a high proportion of households living in thatched roof homes (rather than metal) were prioritized.

Within each village, households were randomly chosen as described in Section 3. Each selected household was then visited by a representative of *GD*. The *GD* representative asked to speak to the member of the household that had been chosen as the transfer recipient *ex ante* (for the purposes of the present study, the recipient was randomly chosen to be either the husband or the wife, with equal probability; details in Section 3). A conversation in private was then requested from this household member, in which they were asked a few questions about demographics, and informed that they had been chosen to receive a cash transfer of KES 25,200 (USD 404). The recipient was informed that this transfer came without strings attached, that they were free to spend it however they chose, and that the transfer was a one-time transfer and would not be repeated.

Recipients were also informed about the timing of this transfer; for the purposes of the present study, 50 percent of recipients were told that they would receive the transfer as one lump-sum payment, and the remaining 50 percent were told that they would receive the transfer as a stream of nine monthly installments. The timing of the transfer delivery was also announced. In the case of monthly transfers, the first installment was transferred on the first of the month following the initial visit, and continued for eight months thereafter. In the case of lump-sum transfers, a month was randomly chosen among the nine months following the date of the initial visit.

For receipt of the transfer, recipients were provided with a SIM card by Kenya's largest mobile service provider, *Safaricom*, and asked to activate it and register for *Safaricom's* mobile money service *M-Pesa* (Jack and Suri 2013). *M-Pesa* is, in essence, a bank account on the SIM card, protected by a four-digit PIN code, and enables the holder to send and receive money to and from other *M-Pesa* clients. Prior to receiving any transfer, recipients were required to register for *M-Pesa*. For lump sum recipients, a small initial transfer of KES 1,200 was sent on the first of the month following the initial *GD* visit as an incentive for prompt registration. Registration had to occur in the name of the designated transfer recipient, rather than any other person. The *M-Pesa* system allows *GD* to observe the name in which the account is registered in advance of the transfer, and transfers were not sent unless the registered name had been confirmed to match the intended recipient within the household. In our sample, all but 18 treatment households complied with these instructions. To avoid biasing our treatment effect estimates, we use a conservative intent-to-treat approach and include data from these 18 non-compliant households in the treatment group.[3] Transfers commenced on the first of the month following registration. Each transfer was announced with a text message to the recipient's SIM delivered through the *M-Pesa* system. However, receipt of these text messages was not necessary to ensure the receipt of transfers; recipients who did not own cell phones could rely on the information about 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 easier communication with recipients and reliable transfer

[3]In a few additional cases, delays in registration occurred due to delays in obtaining an official identification card, which is a prerequisite for registering with *M-Pesa*.

delivery, *GD* offered to sell cell phones to recipient households which did not own one (by reducing the future transfer by the cost of the phone).

Withdrawals and deposits can be made at any *M-Pesa* agent, of which *Safaricom* operates about 11,000 throughout Kenya. Typically an *M-Pesa* agent is a shopkeeper in the recipients' village or the nearest town (other types of businesses that operate as *M-Pesa* agents are petrol stations, supermarkets, courier companies, "cyber" cafes, retail outlets, and banks). *GD* estimates the average travel time and cost from recipient households to the nearest *M-Pesa* agent at 42 minutes and USD 0.64. Withdrawals incur costs between 27 percent for USD 2 withdrawals and 0.06 percent for USD 800 withdrawals, with a gradual decrease of the percentage for intermediate amounts.[4] *GD* reports that recipients typically withdraw the entire balance of the transfer upon receipt.

The sender also incurs costs for *M-Pesa* transfers; according to *GD*'s estimates, the costs of transferring money to recipients in this fashion amount to 1.5 percent of the transfer amount for foreign exchange fees, and 1.6 percent for *M-Pesa* fees. Together with 4.8 percent of transfers spent on recipient identification and staff costs, GD estimates that 92.1 percent of the donations it receives are transferred to recipients' *M-Pesa* accounts.

3 Design and methods

3.1 Experimental design

3.1.1 Sample selection

This study employs a two-level cluster-randomized controlled trial. An overview of the design is shown in Figure 1. In collaboration with *GD*, we identified 126 villages from a list of villages in Rarieda district of Western Kenya. In the first stage of randomization, 63 of these villages were randomly chosen to be treatment villages. Within all villages, we conducted a census with the support of the village elder, which identified all eligible households within the village. As described above, eligibility was based on living in a house with a thatch roof. Control villages were only surveyed at endline; in these villages, we sampled 432 households from among eligible households, to which we refer as "pure control" households in the following.

In treatment villages, we performed a second stage of randomization, in which we randomly assigned 50 percent of the eligible households in each treatment village to the treatment condition, and 50 percent to the control condition. This process resulted in 503 treatment households and

[4] As a result of the *Kenyan Finance Act* of 2012, which introduced a 10 percent excise duty tax on transaction fees for all money transfer services provided by cellular phone providers, banks, money transfer agencies and other financial service providers, *Safaricom* revised the cost structure for sending and receiving money through *M-Pesa*. The costs for transfers over USD 2 increased by 10 percent, while fees remained the same for smaller transfers (http://www.safaricom.co.ke/personal/m-pesa/m-pesa-services-tariffs/tariffs/tariff-faqs). However, these changes did not take effect until February 8, 2013, by which time the endline survey for this study had already been concluded. Our results are therefore unlikely to be affected by this new cost structure.

505 control households in treatment villages at baseline. In the following, we refer to the control households in treatment villages as "spillover" households because comparing these households to control households in treatment villages allows us to identify spillover effects.

As described above, due primarily to registration issues with *M-Pesa*, 18 treatment households had not received transfers at the time of the endline, thus only 485 of the treatment households had in fact received transfers. In the analysis below we use 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.

Due to the fact that the pure control households were selected into the sample just before the endline, the thatched roof criterion was applied to them 12 months later than to households in treatment villages. This fact potentially introduces bias into the comparison of households in treatment and control villages; in the absence of transfers, a proportion of households in treatment villages that had a thatched roof at baseline might have purchased a metal roof independently of the transfers and thus may not be comparable to homes with thatched roofs in pure control villages at endline. However, we can obtain bounds for this potential source of bias by comparing the treatment effects for households that still had thatched roofs at endline to those for the entire sample (since for the former group of households, the bias would go in the opposite direction). We report the results from this exercise in the Online Appendix; it reveals no differences in the results, both qualitatively and quantitatively, and therefore both the within- and across-village treatment estimates are valid. In what follows we focus on the within-village treatment effect; in the presence of positive spillovers, this is a lower-bound estimate of the treatment effect. The across-village treatment effects are reported int he Online Appendix. [5]

To obtain a lower-bound estimate for spillover effects, we compare households which still have thatched roofs at endline to pure control households which still have thatched roofs at endline. The logic behind this choice is the following. First, note that in the absence of spillover effects on roof purchases, this comparison provides an unbiased estimate of the spillover effects for this group of households. Second, relax the assumption of no spillovers and assume instead (as is likely) that spillover effects predominantly induce the better-off control households in treatment villages to upgrade to a metal roof. If this is the case, restricting the sample to households which still have a thatched roof at endline selects for poorer households in treatment villages, but not pure control villages, and thus provides a lower bound estimate of the spillover effect. To be conservative, in what follows we report this lower-bound estimate.

[5]Note that this strategy would overestimate the treatment effect in the presence of negative spillovers. However, we find little evidence for negative spillovers, as discussed below; this includes psychological well-being, i.e. untreated households in treatment villages did not experience a decrease in psychological well-being. The within-village treatment effects therefore provide a conservative estimate.

3.1.2 Treatment arms

A goal of this study was to assess the relative welfare impacts of three design features of unconditional cash transfers: the gender of the transfer recipient; the temporal structure of the transfers (monthly vs. lump-sum transfers); and the magnitude of the transfer. The intervention was therefore structured as follows:

1. **Transfers to the woman vs. the man in the household.** Among households with both a primary female and primary male member, we stratified on recipient gender and randomly assigned the woman or the man to be the transfer recipient in an equal number of households. A further 110 households had a single household head and were therefore not considered in the randomization of recipient gender.

2. **Lump-sum transfers vs. monthly installments.** Across all treatment households, we randomly assigned the transfer to be delivered either as a lump-sum amount, or as a series of nine monthly installments. Specifically, 258 of the 503 treatment households were assigned to the monthly condition, and 245 to the lump-sum condition. The total amount of each type of transfer was KES 25,200 (USD 404). This amount includes an initial transfer of KES 1,200 (USD 19) to incentivize *M-Pesa* registration, followed by either a lump-sum payment of KES 24,000 (USD 384) in the lump-sum condition, or a sequence of nine monthly transfers of KES 2,800 (USD 45) each in the monthly condition. The timing of transfers was structured as follows: in both the monthly and the lump-sum condition, recipients received the initial transfer of KES 1,200 immediately following the announcement visit by *GD*. In the monthly condition, recipients then 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 lump-sum transfer of KES 24,000 on the first of a month that was chosen randomly among the nine months following the time at which they were enrolled in the *GD* program.

3. **Large vs. small transfers.** Finally, a third pair of treatment arms was created to study the relative impact of large compared to small transfers. To this end, 137 households in the treatment group were randomly chosen and informed in January 2012 that they would receive an additional transfer of KES 70,000 (USD 1,112), paid in seven monthly installments of KES 10,000 (USD 160) 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[6], and therefore the total transfer amount received by these households was KES 95,200 (USD 1,525). The remaining 366 treatment households constitute the "small" transfer group, and received transfers totaling KES 25,200 (USD 404) per household.

[6] 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.

These three treatment arms were fully cross-randomized, except that, as noted above, the "large" transfers were made to existing recipients of KES 25,200 transfers in the form of a KES 70,000 top-up that was delivered as a stream of payments after respondents had already been told that they would receive KES 25,200 transfers. Section 5 outlines how this issue is dealt with in the analysis.

3.1.3 Timeline

The timeline of the study is summarized in Figure 2. Baseline surveys took place between May and November 2011, and endline surveys between September and December 2012. Transfers were made between June 2011 and January 2013[7]. Monthly transfers were made in nine monthly installments of KES 2,800 (USD 45), and lump sum transfers were made all at once, in a randomly selected bin among nine monthly bins. Thus, the transfers were timed so that the total amount of lump sum transfers in a given month was the same as the total amount of monthly transfers in that month. For the large transfer group, an additional transfer of KES 70,000 (USD 1,121) was issued in seven monthly installments.

3.2 Data collection

In treatment villages, we surveyed treatment and control households both at baseline and endline; in control villages, we surveyed "pure control" households at endline only. In each surveyed household, we collected two distinct modules: a household module, which collected information about assets, consumption, income, food security, health, and education, administered to either the primary male or female member of the household; and an individual module, which collected information about psychological well-being, intrahousehold bargaining and domestic violence, and preferences. Finally, we measured the height, weight, and upper-arm circumference of the children under five years who lived in the household. The two surveys were administered on different (usually subsequent) 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, i.e. 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, in particular the spouse.[8]

[7] Despite the overlap between the baseline and endline survey periods with the period in which transfers were sent, for each individual recipient the baseline survey was administered before receipt of the first transfer. Nine lump-sum recipient households received their transfers after the administration of the endline survey; data from these households is nevertheless included in the analysis to be conservative.

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

From a randomly selected subset of on average three respondents in each village, we also obtained village-level information about prices, wages, and crime, to assess possible market equilibrium effects of the intervention (Angelucci and De Giorgi 2009). All questionnaires are available from the authors upon request.

An important feature of this study is that, in addition to questionnaire measures of psychological well-being, we also obtained saliva samples from all respondents, which were assayed for the stress hormone cortisol. 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)[9], and it has several practical advantages which make it attractive as an analyze in field studies.[10] 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.[11]

We obtained two saliva samples from each respondent, at the beginning at at the end of the individual survey, 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). 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 non-invasive nature of this technique, we encountered no apprehension among respondents. The saliva samples were labeled with barcodes and stored in a freezer at $-20\,\deg$ C, and were later centrifuged and assayed for salivary free cortisol using a standard radio-immunoassay (RIA) on the *cobas e411* platform at *Lancet Labs*, Nairobi.

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

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

[9] 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, two factors make it unlikely that our respondents undertook such manipulation: first, for this manipulation to systematically affect our results, our participants would have to have intimate knowledge of the environmental and physiological factors that affect cortisol levels; 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 common factors that affect cortisol levels; or participants would have to systematically lie about certain control variables. Given the fact that our respondents largely appeared unaware 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 can be measured non-invasively 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.

[11] 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, however, 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, e.g. in atherosclerosis and cancer (Steptoe et al. 2001; Steptoe et al. 2002; Aggarwal et al. 2006; Coussens and Werb 2002; Ross 1999). Thus, permanently high cortisol is physiologically damaging, over and above its psychological significance.

Cortisol levels were analyzed as specified in our pre-analysis plan, and briefly summarized here: 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-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 are nearly identical when including vs. omitting these control variables.

3.3 Integrity of experiment

3.3.1 Baseline balance

We test for baseline differences between treatment and control groups using the following specification:

$$y_{vhiB} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \varepsilon_{vhiB} \tag{2}$$

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). The sample is restricted to treatment and control households in treatment villages, as explained above. Village-level fixed effects are captured by α_v. T_{hv} is a treatment indicator that takes value 1 for treatment households, and 0 otherwise. ε_{vhiB} is an idiosyncratic error term. The omitted category is control households in treatment villages; thus, β_1 identifies the difference in baseline outcomes between treated households and control households in treatment villages. Standard errors are clustered at the level of the unit of randomization, i.e. the household. In addition to this standard inference, we compute FWER-corrected p-values across the set of index variables. Finally, we estimate the system of equations jointly using seemingly unrelated regression (SUR), which allows us to perform Wald tests of joint significance of the treatment coefficient.

The results of this estimation for our index variables are shown in the Online Appendix . The results are largely insignificant, suggesting that the treatment and control groups did not differ at baseline.

3.3.2 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 had in fact received transfers. 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.

3.3.3 Attrition

We find low levels of attrition overall; overall, 940 of 1,008 (93.3 percent) baseline households could be surveyed at endline. In the treatment group, 471 of 503 baseline respondents (94 percent) were surveyed at endline, and in the spillover group, 469 of 505 (93 percent). Detailed attrition analyses are shown in the Online Appendix. 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. Second, a regression of our main index variables on the attrition dummy reveals no significant overall difference in outcomes at baseline between attrition and non-attrition households. Finally, 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 vs. the control conditions. Thus, attrition is unlikely to have biased the results reported below.

3.3.4 Effects of *M-Pesa* access

A possible concern in the present study is that recipient households were provided with a SIM card and required to register for *M-Pesa*, while control households did not receive a SIM card, nor were they encouraged or required to sign up for *M-Pesa*. In light of recent evidence on the risk smoothing effects of having access to M-Pesa (Jack and Suri 2013), and that simply providing access to a savings device such as M-Pesa can substantially affect household savings (Dupas and Robinson 2013b), this difference between our treatment and control groups raises the question whether the economic effects we observe are a result of *M-Pesa* access *per se*, rather than receipt of cash transfers.

Our data affords us an opportunity to assess this possibility as it contains detailed data on remittances and savings that distinguish whether these were effected using *M-Pesa*. In the Online Appendix, we report results from a regression of different *M-Pesa* use variables on treatment. We find no effect of treatment on the sending of remittances through M-Pesa, but significant treatment effects on receiving remittances through *M-Pesa* and saving with *M-Pesa*. However, two facts argue against the possibility that this increase in the use of *M-Pesa* was a main driver of our effects.

First, the magnitude of the effects is small compared to the size of our treatment estimates on outcomes such as assets, consumption, and agricultural and business income. Specifically, treatment households save an extra USD 3 in M-Pesa compared to control households, and receive an extra USD 9 per month in remittances. Both of these effects are relatively small compared to e.g. the USD 36 monthly increase in consumption among treatment households.

Second, Aker et al. (2011) conducted an RCT in Niger in which they compared a group which received a manually delivered cash transfer (USD 45 per month) to a second group which received the same cash transfer, but additionally was given access to the mobile money system *ZAP*. This study this tests precisely the additional effect of having access to a mobile money system over and above that of a cash transfer. Across a wide variety of outcome variables including assets, food security, consumption, income and agricultural production, migration, and remittances, no significant differences were found between these two groups. Thus, it appears, at least in this setting, that simply providing access to a mobile money technology cannot change household outcomes significantly, over and above the effects of a large cash transfer.

3.4 Data analysis

3.4.1 Pre-analysis plan

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 paper to report the full breadth of the evidence. However, ambition makes it necessary to find a way to report all relevant results without being overwhelming and without cherry-picking. To discipline our analysis of the reduced from evidence, we wrote a pre-analysis plan (PAP) for this study, which is published and time-stamped at www.socialscienceregistry.org (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 indices, 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 paper 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. However, this restriction is conservative. The PAP did not cover the extra analysis that is described in Sections 5.1 and 5.2 (calculation of elasticities and discount rates); however, these analyses follow standard methods, and in addition their goal is not to estimate treatment effects, but to understand mechanisms.

3.4.2 Reduced form specifications

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

$$y_{vhi} = \alpha_v + \beta_0 + \beta_1 T_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhi} \tag{3}$$

where y_{vhi} 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). The sample is restricted to treatment and control households in treatment villages, as explained above. Village-level fixed effects are captured by α_v. T_{hv} is a treatment indicator that takes value 1 for treatment households, and 0 otherwise. ε_{vhi} is an idiosyncratic error term. The omitted category is control households in treatment villages; thus, β_1 identifies the treatment effect for treated households relative to control households in treatment villages. 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 zero and include a dummy indicator that the variable is missing (M_{vhiB}). Standard errors are clustered at the level of the unit of randomization, i.e. the household. In addition to this standard inference, we compute FWER-corrected p-values across the set of index variables. Finally, we estimate the system of equations jointly using seemingly unrelated regression (SUR), which allows us to perform Wald tests of joint significance of the treatment coefficient across outcome variables.

To distinguish between the effects of different treatment arms, we proceed as follows. First, the effect of making the transfer to the female vs. the male in the household is captured by the following model:

$$y_{vhi} = \alpha_v + \beta_0 + \beta_1 T^{\mathrm{F}}_{vh} + \beta_2 T^{\mathrm{M}}_{vh} + \beta_3 T^{\mathrm{W}}_{vh} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhi} \tag{4}$$

In this specification, the sample is restricted to cohabitating households in treatment villages. The variables T^{x}_{vh} are indicator functions that specify the branch of the different treatment arms. Specifically, they indicate whether the transfer recipient is female (T^{F}_{vh}), male (T^{M}_{vh}), or that the gender of the recipient could not be randomized because the household only had one head (T^{W}_{vh}; most commonly in the case of widows/widowers). The coefficients on T^{F}_{vh} and T^{M}_{vh} identify the treatment effect for female recipient and male recipient households, respectively, relative to control households in treatment villages. The significance of the difference is estimated with a Wald test. Village fixed effects are captured by α_v, and standard errors are clustered at the household level. Note that spillover effects cannot be identified separately for the sub-treatments since these were randomized within villages and there was little variation in the intensity of each arm of a given sub-treatment across villages.

To asses the relative effect of monthly vs. lump-sum transfers, recall first that a subset of households originally assigned to receive USD 404 in either lump-sum or monthly transfers was randomly assigned to receive additional monthly transfers beginning in February 2012 to achieve a total transfer of USD 1,525. Households in this category which had previously been assigned the lump-sum condition can therefore not be unambiguously assigned to either the lump-sum or monthly condition. To control for this ambiguity, the regression comparing lump-sum and monthly transfers is estimated only for the groups which did not receive the large transfers. The specification is as follows:

$$y_{vhi} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\mathrm{MTH}} + \beta_2 T_{vh}^{\mathrm{LS}} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhi} \tag{5}$$

In this specification, T_{vh}^{MTH} and T_{vh}^{LS} are indicator variables for having *originally* been assigned to receive monthly or lump-sum transfers, respectively. The sample is restricted to households in treatment villages, and as explained above, only small transfer recipients are included in the treatment group.

Finally, to assess the effect of receiving large compared to small transfers, we use the following specification:

$$y_{vhi} = \alpha_v + \beta_0 + \beta_1 T_{vh}^{\mathrm{L}} + \beta_2 T_{vh}^{\mathrm{S}} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhi} \tag{6}$$

Again the sample is restricted to households in treatment villages. T_{vh}^{S} and T_{vh}^{L} are indicators for being assigned to the small vs. large transfer condition, respectively.

Spillover effects are identified as follows (Duflo and Saez 2003):

$$y_{vhi} = \beta_0 + \beta_1 S_{vh} + \varepsilon_{vhi} \tag{7}$$

Here, the sample includes only non treatment households (in treatment and control villages), and S_{vh} is a dummy variable that takes value 1 for control households in treatment villages ("spillover households") and 0 otherwise. In this specification, we restrict to households with a thatched roof at endline to account for the delay in which the selection criterion was applied to the pure control village (discussed above). Thus, β_1 identifies within-village spillover effects by comparing control households in treatment villages to control households in pure control villages. The error term is clustered at the village level, reflecting the dual-level randomization at the village and within-village (household) level (Cameron, Gelbach, and Miller 2011; Pepper 2002). Note that the inclusion of baseline covariates is not feasible here because no baseline data exists for the pure control group. Similarly, village-level fixed effects are not feasible because they would be collinear with S_{vh}.

3.4.3 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 et al. (2007), Anderson (2008), and Casey et al. (2012).

First, we compute standardized indices for several main groups of outcomes, and choose focal variables of interest for others (all specified in the PAP). In particular, we use the total value of household assets, total household consumption in the past month, and total household agricultural and business income in the past month, as focal variables for the asset, consumption, and income outcome groups, respectively. For psychological well-being, food security, female empowerment, health, and education, we compute indices, which are standardized weighted averages of several key outcomes of interest within each of these groups of outcomes. The particular outcomes composing each index and the focal variables were pre-specified in our pre-analysis plan.

Second, even after collapsing variables into indices and choosing focal variables of interest for each group of outcomes, we are still left with multiple indices, creating the need to further control the probability of Type I errors. To this end, we use the Family-Wise Error Rate (FWER; Westfall, Young, and Wright 1993; Efron and Tibshirani 1993; Anderson 2008; Casey, Glennerster, and Miguel 2012), which controls the probability of Type I errors across a group of coefficients. In our case, we control the FWER across the treatment coefficients on the indices for our main outcome groups, i.e. assets, consumption, income, psychological well-being, education, food security, health, and female empowerment. As specified in our pre-analysis plan, 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 to single out particular variables for general conclusions.

4 The impact of unconditional cash transfers: A summary of the reduced form evidence

In this section, we present an overview of the reduced form effect of the transfers. A fuller description of the results is available in a Policy Brief and the Online Appendix, both at web.mit.edu/joha/www.

Table 1 shows the main results of the program for the index variables and key outcomes pre-selected in the pre-analysis plan. Column (1) shows the mean and the standard deviations in the control group. The treatment and spillover effects for these variables are shown in columns (2) and (3), estimated using equations 3 and 7. The subsequent columns show the differences between the individual treatment arms. Standard errors are shown in parentheses, and the bootstrapped FWER p-values in brackets. The last row of the table reports the joint significance of all coefficients in the corresponding column, using seemingly unrelated regression (SUR).

4.1 Direct effects on beneficiaries: overall effects and impact on psychological wellbeing

Overall impacts

On average, households were surveyed 4.3 months after receiving their last transfer. We find statistically and economically meaningful impacts of cash transfers across the majority of outcomes measured by our indices, 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 0.001. Household consumption of nondurable goods is significantly higher in the treatment group than in the control group (USD 36 per month, or 23 percent of the control group consumption). The effect is statistically significant at the 1 percent confidence level according to both standard and FWER p-values. Concomitantly, we see improvement in the food security index (0.25 SD) – we will return to these results in the next section, where we investigate the expenditure elasticity of various budget heads. Households invest part of the transfers: the value of non-land assets increased by USD 279 on average; this represents 58 percent of the control group mean USD 478, and 39 percent of the average transfer.[12] The effect is statistically different from zero at the 1 percent 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 USD 15 increase. This is an increase of 33 percent over the control group mean, and on an annual basis, it represents 26 percent of the average transfer amount. We see no improvement in health, education, or female empowerment when comparing treatment and control households within the same villages. However, the female empowerment result is accounted for by gains among spillover households, which we document below.

Effects of treatment arms

We now discuss each of the three sub-treatments in turn: transfers to the primary female vs. the primary male in the household; monthly vs. lump-sum transfers; and large vs. small transfers.

Column (4) in Table 1 reports the coefficients and standard errors comparing female to male recipient households on the index variables. With the exception of psychological well-being, which is significant at the 10 percent level and further discussed below, none of the differences between the treatment effects for transfers to the female vs. the male are statistically significant at conventional significance levels. Thus, we find little evidence that providing cash transfers to women vs. men differentially affect outcomes. However, we note 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

[12] 28 percent of the treatment group received a transfer of KES 95,200 (USD 1,525), while the remaining 72 percent received KES 25,200 (USD 404); the average transfer was thus KES 44,800 (USD 717).

female in the household improves outcomes most likely to benefit children, i.e. food security, health, and education, as well as psychological well-being and female empowerment.

Results comparing monthly to lump-sum transfers are shown in column (5) of Table 1. The joint significance across outcomes is at $p < 0.05$, 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 SD relative to lump-sum payments, and that lump-sum transfers lead to higher asset values than monthly transfers; both effects are statistically significant at conventional levels, but not using FWER-corrected standard errors.

Finally, column (6) of Table 1 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 outcome has a p-value of less than 0.001. 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 percent level using conventional p-values, but not FWER-corrected inference.

Psychological well-being and cortisol

A central goal of this study was to assess in detail the effects of unconditional cash transfers on psychological well-being and levels of the stress hormone cortisol. Based on our previous work (Chemin et al. 2013; Haushofer et al. 2011; Haushofer and Fehr 2013), we had hypothesized that cash transfers would lead to an increase in psychological well-being, and specifically to a reduction in stress and cortisol levels. Overall, the transfers indeed led to a large and significant improvement in psychological well being of 0.20 SD, significant at the 5 percent level according to both standard and FWER adjusted p-values. Table 7 investigates this effect in more detail, and shows that it stems mainly from a 0.18 SD increase in happiness scores, a 0.15 SD increase in life satisfaction, a 0.14 SD reduction in stress, and a 0.99 point reduction in scores of the CESD depression questionnaire[13] (see Online Appendix for details on the variables). Thus, we indeed observe significant reductions in stress and depression, and increases in happiness and life satisfaction, as a result of exogenous reductions in poverty, lending support to our hypothesis that poverty alleviation may have psychological benefits.

In addition, while the average effect of treatment on cortisol levels is small and not significant, we find large and significant differences in cortisol levels between all treatment arms: first, men

[13]The control group mean of this questionnaire was 26.5. The established cutoff for the presence of depression is a score of 16.

and women in households where the female received the transfer had significantly lower cortisol levels than respondents in households where the male received the transfer. The magnitude of the effect is 0.21 log units, which corresponds to a difference between female and male recipient households of 2.73 nmol/l. This is a large effect when compared against, for example, 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. In addition, we 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 a) psychological well-being correlates highly with female empowerment in the cross-section (Online Appendix), and that b) female empowerment shows a relatively large, although non-significant, difference of 0.16 SD between male and female recipient households (in fact, this difference is of the same magnitude, 0.16 SD, as the difference in psychological well-being between male and female recipient households). 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. 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).

Second, we find a large and significant difference in cortisol levels between monthly and lump-sum recipient households: in monthly recipient households, cortisol levels are 0.27 log units (3.63 nmol/l) higher than in households that receive their transfers as a lump sum. 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 lump-sum transfers, and food security correlates well with psychological well-being in the cross-section, we would 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 to save the transfers; thus, it is possible that the increased cortisol levels in this condition reflect the stress of being unable to save or invest the transfers, despite better intentions.

Finally, we find that cortisol levels are 2.03 nmol/l lower in households that received large transfers than in households that received small transfers. Concomitantly, we observe very large additional gains in psychological well-being for large transfers: the overall index of psychological well-being for the large transfer group is a full 0.45 SD higher in the treatment than in the control group.

This effect is very large both in absolute terms, and relative to the effect of the small transfer (0.11 SD). Apart from cortisol, this effect is driven by a large difference in depression scores of 1.76 points between large and small transfer recipients, a 0.30 SD difference in stress scores, and 0.17 SD difference in life satisfaction.

Together, these findings provide additional support for our ingoing hypothesis that poverty alleviation would lead to improvements in psychological well-being and decreases in cortisol levels. More broadly, they 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.

4.2 Spillovers to other households and equilibrium impacts

Column (3) in Table 1 reports the coefficients on the spillover dummies. These are generally small and not significant, with one exception: we observe an increase of 0.23 SD in the female empowerment index among the control group in treatment villages. This increase is significant at the 5 percent level using conventional p-values. Together with a non-significant direct treatment effect of SD -0.01 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 the Online Appendix. However, since this is the only outcome that shows any spillover effect and 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, and instead note that it needs to be replicated.

Two variables are particularly interesting. First, we find no spillovers in consumption. This is surprising, given that we might have expected some informal insurance among households: in effect, the transfer is a temporary lottery gain, and theory predicts that households should have been sharing it with their insurance network (Townsend 1994). This finding also contrasts with (non-experimental) evidence from conditional cash transfer programs (Angelucci and De Giorgi 2009).

A second interesting variable is the potential spillover on psychological well-being among non-beneficiaries. One potential concern about delivering cash transfers to some households in a village but not others is the potential for negative externalities: the untreated households may experience a reduction in psychological well-being simply by virtue of not having received transfers. This does not seem to have occured here. If anything, spillover effects on psychological well-being were broadly positive: most coefficients on the spillover dummies go in the direction of "positive" outcomes, although they are not significant overall.[14]

The lack of spillovers to non-treatment households suggests that the transfers had no major impact on the economic environment. To investigate more directly whether individual cash transfers caused

[14]There is a marginally significant impact on optimism, which may reflect the fact that control households were expecting to receive transfers, from *GiveDirectly* or elsewhere.

market equilibrium shifts (in 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 indices as described in the Online Appendix. We regress average village-level outcomes ($\bar{y}$) on an indicator variable for whether village v is a treatment village:

$$\bar{y}_v = \beta_0 + \beta_1 T_v + \varepsilon_v \tag{8}$$

We present results in the Online Appendix. We find no significant village-level effects on any variable group, and thus conclude that the transfer 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[15].

5 How households spend their money

In this section, we exploit both the program and the fact that the experimental design purposefully varied key features to shed more general light on households response to a temporary (but large) change in income, and thereby address several long-standing questions in the literature. Specifically, we estimate the expenditure elasticities of a broad set of budget heads and compare them to cross-sectional estimates; further, comparing elasticities across different treatment arms allows us to assess to what extent i. households are unitary, ii. households are credit- and savings-constrained, and iii. Engel curves are concave.

5.1 Expenditure elasticities: an instrumental variable approach

Assessing the increase in permanent income caused by the program would require a full understanding of households' savings, borrowing and investment behavior and opportunities, and knowledge of what they expected regarding future transfers, and is beyond the scope of this paper. However, when households were surveyed, on average 4.3 months after the end of the program, we observe an increase in monthly flow of non-durable expenditure of USD 36 relative to a control group mean of USD 157. Under standard time-separability assumptions, we can use the transfer as an instrument for the change in non-durable expenditure and estimate the elasticities of different expenditure heads with respect to total non-durable expenditure. This is a classic question in the economics of consumption (Deaton 1992). In particular, the elasticity of food expenditure with respect to overall expenditures is a key parameter both for many economic models, including the Dasgupta & Ray (1986) model of poverty traps, and for policy design. Estimates vary greatly, depending on different

[15] Another possible explanation for these null finding is that effects at the level of the local economy are larger for in-kind transfers than cash transfers (Cunha, De Giorgi, and Jayachandran 2011).

empirical strategies, and there is little consensus on the size of the elasticities (Strauss and Thomas 1995).

Table 2 shows detailed results on the impact of the transfer of non-durable expenditure variables. With the exception of temptation goods (defined as spending on alcohol, tobacco and gambling), cash transfers increase all categories of consumption, including food, medical and education expenses, durables, home improvement, and social events. On an absolute basis, the largest increases in consumption are food (USD 20 per month, a 19 percent increase). Spending on medical care, education and social expenditures (e.g., weddings, funerals, recreation) increase significantly on a percentage basis, but from relatively lower levels. Spending on other items, including airtime, household and personal goods increases by USD 10. These impacts amount to a total increase of USD 36 in non-durable expenditures among the treatment group.

We next estimate the elasticities of expenditure in various categories with respect to total non-durable expenditures that are implied by these numbers. To compare with the observational data, Figure 4 shows local linear estimates for cross-sectional Engel curves for different expenditure heads; the red line represents the treatment group, the blue line the control group. Because the axes are log-log, the slope corresponds to the elasticity. We then superimpose the average changes in overall expenditure and food expenditure induced by treatment; the large blue dot represents the log of mean monthly total and food consumption in the control group at endline, and the red dot in the treatment group.

Since the relationship appears to be relatively linear, we first estimate the following cross-sectional model (Deaton and Subramanian 1996).[16]

$$\ln\left(x_{vhj}\right) = \alpha_v + \beta_0 + \beta_1 \ln\left(y_{hv}\right) + \varepsilon_{hv} \tag{9}$$

Here, x_{vhj} is expenditure on budget head j in household h in village v at endline, and y_{hv} is total endline expenditure. The sample is restricted to the spillover group.

There is a large literature estimating such regressions, but the cross-sectional estimates potentially conflate actual elasticities with underlying differences between richer and poorer households. Indeed, several recent pieces of evidence have suggested that cross sectional Engel curves could be biased upwards. In India, Deaton & Dreze (2002) find that they have been sliding over time. In China, Jensen and Miller (2008b) find that calorie availability (although not food expenditure) actually significantly declined with a subsidies on staples (with is only possible if income effects on calories are strong and negative). Together with the large variability in elasticities reported across studies (1995), these results raise the question whether experimentally induced variation in expenditures would produce different estimates.

[16]To deal with zeroes in expenditure for some items, we again use the inverse hyperbolic sine transform described above (Burbidge, 1988; MacKinnon, 1990; Pence, 2006).

To address this question, we estimate a version of Equation 9 in which we instrument total expenditure with a dummy for being in any of the treatment groups. In this specification, the sample includes both the treatment and within-village control group. We then present and compare both OLS estimates and IV estimates.

In Table 3 , columns (1) and (2) shows the OLS and IV estimates respectively for different expenditure and consumption categories for the entire sample. The Hausman tests for equality of the OLS and IV estimates are shown in column (3). The food expenditure elasticity (food expenditures include the consumption of food produced at home) is strikingly large: the IV estimate is 0.83, larger than the Subramanian & Deaton estimate of 0.62. The OLS estimate is even larger (1.00). The difference between the OLS and IV estimate is significant, but both are high (perhaps surprisingly so), and in the same range. The IV estimate suggests that, in this sample, the amount spent on food increases almost proportionally with income.

We hasten to point out that the food expenditure elasticity is not the calorie elasticity: as households spend more on food, they also buy more expensive calories (in the Subramanian and Deaton estimate, for every extra rupee spent on food, 0.5 rupees are spent on more calories, and 0.5 on more expensive calories). Unfortunately, we did not collect detailed quantity data for food items and therefore cannot compute calorie elasticities. However, we clearly observe some substitution towards more expensive food items: the elasticity is highest for meat and fish and dairy (> 2), and lowest for cereals and sugar (it is possible that sugar has a relatively low elasticity because there is less variation in quality within sugar). However, the elasticities are large for all food items, making it likely that both quantity and quality increased.

This conjecture is confirmed in Table 4, where we study the treatment effects on food security. To increase sensitivity, where possible we use as outcome variables the *number* of instances in the past month where households suffered from different types of food insecurity, such as having to skip meals, going for whole days without food, having to rely on others for help with food, etc. Four variables are only available as dummies (going to bed hungry, regularly eating two meals, eating until content every day, and having enough food for the next day).

Food security is low in this population: 42 percent of households report that adults have gone without food for an entire day in the past month, 86 percent of households have relied on help from others for food, and 36 percent had to beg due to a lack of food. With one exception (where the coefficient is small and non-significant), all of the signs of the treatment coefficients are in the direction of increased food security, resulting in a statistically significant 0.11 SD increase in our food security index. This effect is driven by a broad range of individual variables, many of which are individually highly significant. For instance, cash transfers reduce the likelihood of the respondent having gone to bed hungry in the preceding week from 22 percent to 14 percent (a 36 percent decrease), increase the likelihood of having enough food in the house for the next day from 35 percent to 42 percent (20 percent), and reduce by 42 percent the number of days children go without food.

Turning to other elasticities, the IV elasticities are greater than unity for medical and social expenditure (which are luxury goods), below one but still fairly large for education expenditure, and negative and not significantly different from zero for alcohol and tobacco.

A 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 instead 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 only with 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. Table 2 in the Online Appendix suggests not only that there was not treatment effect on alcohol and tobacco consumption when using this method, but additionally shows 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 injects noise into the data, and the results are therefore imprecise.

Are the elasticities constant for different types of income changes?

There are several reasons why the elasticities of different expenditure categories with respect to overall expenditure could vary: households may not be unitary, in which case recipient gender might matter; they may be credit and/or savings constrained, in which case the delivery schedule (monthly vs. lump-sum) could matter; and returns to transfers may be increasing or decreasing, in which case transfer size might matter. For this reason, our experimental design varied the recipient, the time structure, and the magnitude of the transfer. As shown in Table 1, the different types of transfer appear to have had comparable impacts on overall non-durable expenditures, except, quite naturally, for the large transfer. The more detailed reduced form estimates in Table 2 suggest that, for most expenditure categories individually, and overall (see the p-values for the joint test in the last row of the tables), cash transfers had similar impacts on all expenditure types, regardless of transfer recipient and timing, and that larger transfers had larger impacts on most goods.

To formalize these conjectures, we test the hypothesis that elasticities estimated with pairs of instruments (e.g. large vs. small transfers) are the same, using a nonlinear over-identification test (this strategy is similar to Duflo and Udry 2004). To compare the coefficients of different treatment arms in the IV specification against each other, we jointly estimate the first-stage and reduced-form regressions for the desired treatment arms using seemingly unrelated regression. Using large vs. small transfers as an example, the system of equations is as follows:

$$\ln(x_{vhj}) = \alpha_v + \beta_0 + \beta_1 T^L_{vh} + \beta_2 T^S_{vh} + \varepsilon_{vhj} \tag{10}$$

$$\ln(y_{vh}) = \alpha_v + \gamma_0 + \gamma_1 T^L_{vh} + \gamma_2 T^S_{vh} + \eta_{vh} \tag{11}$$

Equation 10 is the reduced form regression of category j expenditure x_{vhj} on the treatment indicators T^L_{vh} and T^S_{vh} , and Equation 11 the first stage for the IV regression of category j expenditure x_{vhj} on total expenditure y_{vhi}, using T^L_{vh} and T^S_{vh} as instruments for y_{vh}. As above, α_v are village-level fixed effects, and ε_{vhj} and η_{vh} are error terms. Analogous systems are estimated to compare monthly and lump-sum transfers, and transfers to the male vs. to the female, each time restricting the samples as described above.

We then test the IV estimators for the different treatment arms implied by these estimates against each other using the delta method. The restriction is as follows:

$$\frac{\beta_1}{\gamma_1} - \frac{\beta_2}{\gamma_2} = 0$$

The results are presented in Table 3; the last row presents tests of joint significant of all elasticity estimates in a given column using SUR. We are now in a position to address the questions posed above about the effects of different types of transfers on expenditure elasticities.

A first important question is whether the Engel curves are in fact linear in logs as in Equation 9, or whether instead the elasticity is declining in total expenditure for some goods (e.g. food, a necessity) and increasing for others (e.g. social expenditures if they are considered a luxury). The distinction between large and small transfers in our design allows us to address this question. We find that over the range of expenditures in our sample, the elasticities are close to constant for all goods but food. The food expenditure elasticity is indeed lower when instrumented with the large transfers, which suggests some concavity, although the p-value for the ratio test is only 0.10. This concavity is driven by dairy and sugar consumption. The point estimates for the elasticity of medical expenditure are markedly different when using small vs. large transfers as instruments, but they are imprecisely estimated and therefore not statistically different. All other elasticities are similar when estimated from large or small transfers.

Second, we compare the expenditure elasticities for lump-sum vs. monthly transfers. While most elasticities are similar, the food expenditure elasticity is larger for monthly than lump-sum transfers. We will return to this result, and a possible explanation, below.

A third important hypothesis we test is whether the elasticity of different expenditure items is different when transfer recipients are female vs. male. A large literature suggests that households may not be unitary (Browning and Chiappori 1998), and a common test of whether households are unitary or not is precisely the test of income pooling, which asks whether expenditure shares are different when money is received by husbands vs. wives. Several papers have found this to be the case in observational data (Thomas 1989; Hoddinott and Haddad 1995; Doss 1996; Lundberg, Pollak, and Wales 1997; Aker et al. 2011; Duflo 2003); in particular, female income is associated with larger expenditures on food and on children. Interestingly, we find here that the food expenditure elasticities are very close for male and female recipients (0.88 vs. 0.77), and the difference between them is not significant. Correspondingly, as we see in Table 4, the food security variables (including the overall index) do not increase differentially for female and male recipient households. The only significant difference in consumption patterns is that transfers to the female lead to a large increase in expenditures on cereals (with an implied elasticity of 1.43), while transfers to the male lead to a small and insignificant decrease. The same is true for medical expenditures, particularly for children. Conversely, the elasticity of social expenditures (parties and festivals) is much larger for transfers to the male than for transfers to the female (2.44 vs 0.71). While these differences go in the expected direction, the relatively muted effects are surprising in view of the literature.

This similarity between female and male recipient households is present both for non-durable expenditure composition and asset accumulation and, as apparent in Column (4) of Table 1, is reflected in the overall outcomes: using FWER corrected p-values, the food security index, as already noted, but also the health, education, psychological well-being, and even female empowerment indices are not significantly different when women receive the transfer compared to when men do. The only significant difference between female vs. male recipients is in psychological well-being, which is not significant with FWER corrected p-values, but is significant at conventional levels. As discussed above, this finding is driven by a large difference in cortisol levels between female and male recipient households: cortisol levels in both men and women are reduced to a greater extent when women receive the transfer compared to when men receive it. In addition, when the woman receives the transfer we also observe a greater reduction in worries and a greater increase in self-esteem compared to when the man receives the transfer.

One possible explanation for the absence of differential effects of recipient gender on most outcomes is that because the transfers were explicitly temporary, they did not affect the bargaining power of the household members. This explanation is consistent with the results of Benhassine et al.

2013, who find no differences in outcomes between male vs. female recipients of conditional and unconditional (though loosely tied to school enrollment) cash transfers in Morocco; however, it contrasts with Duflo and Udry 2004, who find that even transitory exogenous changes in male and female income affect a number of expenditure shares. In addition, this account does not explain why transfers to women would differentially increase psychological well-being compared to transfers to men if they do not change other outcomes. Regardless of whether this explanation is correct, this set of results deserves future research; in our view, it is of particular interest whether the cortisol result replicates in other samples.

5.2 Saving for a rainy day?

Because households were explicitly told that the transfers were transitory, we would expect them to save a significant portion of the transfer, or invest in durable goods or productive activity. Did households in fact do this? And did their investments translate into increases in income which could facilitate a permanent increase in consumption? This section addresses these questions[17].

Assets and investment

We begin by estimating the treatment effect on asset ownership. These results are presented in Table 5. The overall treatment effect on assets amounts to USD 279, and is mainly driven by investment in livestock, furniture and metal roofs. Livestock holdings increase by USD 85, a 51 percent increase relative to the control group mean, and 12 percent of the average transfer. This increase extends to all categories of livestock, with the largest increase in absolute terms occurring in cattle holdings (USD 57 relative to a control group mean of USD 102, i.e. 56 percent). Similarly, the value of durable goods owned by treatment households increased by USD 53 relative to a control group mean of USD 207 (an increase of 26 percent, or 7 percent 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 (USD 10).

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 23 percentage points relative to a control group mean of 16 percent. Because the cost of a metal roof is USD 564, this effect corresponds to a USD 130 increase in the value of roofs owned by treatment households, which in turn corresponds to 18 percent of the average transfer. 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.

[17]Note that households may not have fully believed that transfers were in fact transitory, and we have no robust way to estimate the extent to which they thought it was transitory or permanent; however, the investment patterns we discuss below suggest that they did believe it.

Are these investments productive?

Some of the investments we observe are in household durables such as furniture, which are not productive; however, other investments are in inputs to income-generating activities, such as livestock. Did these investments lead to an increase in income?

Table 6 presents impacts of cash transfers on income-generating activities, in particular agricultural and business activities. In analyzing the effect of transfers on these outcomes, we distinguish between a) non-agricultural enterprises, b) agriculture, and c) animal husbandry. In each case, revenue includes own consumption, and we distinguish between flow costs and fixed costs.

There is little evidence that transfers change the primary source of income for recipient households; they are no more or less likely to report farming, wage labor, or non-agricultural businesses as their primary source of income than control households. Cash transfers increase investment in and revenue from income generating activities, primarily non-agricultural businesses and livestock. Monthly expenditure on nonagricultural enterprises (e.g., inputs and inventory) increases by USD 10, with a corresponding increase of USD 11 in monthly revenues from these activities. Similarly, spending on food and care for livestock increases by USD 1 per month, while revenue from the sale of animal products (e.g., milk and eggs) increases by USD 3, 46 percent higher than the control group average. Recipient households also report USD 2 higher revenue from the sale of livestock and meat than control households. These effects translate into an overall increase in monthly gross revenues for recipient households of USD 17 relative to the control group (including revenues from the sale of animals and meat). We do not observe a significant increase in estimated profits from self-employment; in addition, we note that the profits from these activities, which account for all income-generating activities for most households, are lower than our estimates of monthly consumption, suggesting that this analysis may underestimate profits.

Temporal dynamics

Having established that households invest in productive assets, we now ask whether these investments led to persistent increases in consumption. To this end, we first obtain an estimate of how the treatment effects unfolded over time. Our data includes a moderate degree of temporal variation in the end date of the transfers, enabling us to ask whether the treatment effects outlasted the period during which households received transfers. We stress, however, that the current study was not designed to investigate long-term effects; further endline surveys will be required to obtain a more complete understanding of long-term impacts.

We begin by creating separate indicators for the transfer having been completed a specific number of months before the endline survey. In doing so, we allow the temporal dynamics to vary based on whether the household received a lump or monthly transfer, since we might expect the impacts of the large lump-sum transfers to unfold differently over time than that of the smaller monthly transfers. Further, since there is limited variation in the time since the end of the transfer for households receiving large transfers, we restrict the sample to households receiving a small transfer.

We then create indicator variables for time elapsed since the end of transfers. Specifically, we first define a dummy for households that receive transfers contemporaneously, i.e. within the last month. Second, we perform a median split on the delay since the last transfer for the remaining households, which results in one group of households which received their last transfer one to four months ago, and another group which received their last transfer more than four months ago. We then estimate the following model:

$$\begin{aligned} y_{vhi} =& \beta_0 + \beta_2 T_{vh}^{\text{LS:<1}} + \beta_3 T_{vh}^{LS\text{:1-4}} + \beta_4 T_{vh}^{\text{LS: 4+}} + \beta_5 T_{vh}^{\text{MTH:<1}} + \beta_6 T_{vh}^{\text{MTH: 1-4}} \\ &+ \beta_7 T_{vh}^{\text{MTH: 4+}} + \delta_1 y_{vhiB} + \delta_2 M_{vhiB} + \varepsilon_{vhi} \end{aligned} \tag{12}$$

In this specification, T^{x-y} takes value 1 if the transfer was completed between x and y months prior to the survey. The sample is restricted to small transfer recipient and control households in treatment villages.

In Figure 3, we show the resulting treatment effect estimates separately 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, separately for monthly and lump sum households. We also list the p-values from Wald tests for joint significance at each time horizon across the monthly and lump-sum groups. The figure indicates that the observed average impact on overall asset values in treatment households persists over time: both for households receiving lump-sum and monthly transfers, levels of asset holdings are significantly higher than in the control group at all time horizons. We observe no decrease over time in either group; the 95 percent confidence intervals of the coefficients overlap across all time horizons.

Similarly, consumption is elevated relative to control at all time horizons. The point estimates suggest declining impacts on total non-durable consumption over time for the group receiving monthly transfers, though not for the lump sum group; however, the confidence intervals of the treatment effects at shorter vs. longer time horizons overlap, i.e. these differences are not statistically significant.

For agricultural and business revenue, we find no strong indication of changing impacts over time; however, note that the treatment effects are small overall and not distinguishable from zero in this restricted and highly disaggregated sample.

Cash transfers also had persistent impacts on food security. This effect is driven by the monthly transfer group; in the lump-sum group we find little evidence of treatment effects at any time horizon, consistent with the larger overall treatment effect on food security in the monthly group. We also observe that the impact on food security is largest among the group receiving contemporaneous transfers; the treatment effect on food security falls by more than 50 percent over time in the group receiving monthly transfers, although remains positive and statistically different from zero.

The temporal dynamics of cash transfers for additional outcomes (psychological well being, health, education, female empowerment) reveal no differential impacts at different delays. In the case of the treatment effect on psychological well-being, this reflects the fact that the restricted sample used here is underpowered to detect the overall treatment effect we observe for this outcome measure, and the fact that psychological well-being is driven by the large transfer list, which is excluded here.

Estimating the discount factor

The temporal dynamics of the treatment effect identified in the previous section allow us to perform a simple calibration exercise in which we estimate the discount factor using a standard infinite-horizon utility maximization problem. Specifically, consider a household that maximizes a standard problem of the following form:

$$\max_{\{c_t, s_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u\left(c_t\right)$$

subject to

$$\sum_{k=0}^{\infty} \frac{1}{(1+r)^k} c_{t+k} = \sum_{k=0}^{\infty} \frac{1}{(1+r)^k} y_{t+k}$$

where c_t, y_t, and s_t are per-period consumption, income, and savings, respectively. Maximization is straightforward and yields the familiar Euler equation:

$$u'\left(c_t\right) = \beta\left(1+r\right) u'\left(c_{t+1}\right)$$

As an aside, recall that the permanent income hypothesis (Friedman 1957; Modigliani and Brumberg 1954) states that if $\beta = 1 + r$, a one-time, unanticipated income change increases/decreases consumption in each period by the return on the change.

To obtain an estimate of the discount rate implied by a given rate of return and consumption pattern, assume CRRA utility:

$$u(c_t) = \frac{c_t^{1-\rho}}{1-\rho}$$

This assumption simplifies the Euler equation to:

$$\left(\frac{c_{t+1}}{c_t}\right)^{\rho} = \beta\left(1+r\right)$$

We next calculate estimates for the variables in this expression to derive an estimate of β. First, we obtain values for c_t and c_{t+1} by estimating consumption responses to transfers at different time horizons. The observed evolution of total expenditure among transfer recipients is as follows: in the treatment group we find a total expenditure level of USD 190 between 1-4 months after the end of transfers, and USD 180 for transfers that ended more than 4 months ago. The mean time elapsed between the end of transfers and the endline survey is 2.66 months for the first bin and 7 months for the last bin, i.e. the time elapsed between these two dates is 4.33 months. Second, we use $r = 0$ as the rate of return on transfers in terms of profits from self-employment. Finally, we obtain an estimate of $\rho = 3$ from a risk preference questionnaire that was included in the questionnaire. Together, these values allow us to obtain an estimate of β:

$$\left(\frac{180}{190}\right)^3 = \beta^{4.33}$$
$$\beta_{\text{monthly}} = 0.96$$
$$\implies \beta_{\text{yearly}} = 0.64$$

For the discount rate δ, this implies:

$$\delta_{\text{monthly}} = \frac{1 - \beta_{\text{monthly}}}{\beta_{\text{monthly}}} = 0.04$$
$$\delta_{\text{yearly}} = \frac{1 - \beta_{\text{yearly}}}{\beta_{\text{yearly}}} = 0.56$$

Thus, the fairly steep drop in non durable consumption by households suggests a fairly high degree of impatience. Consumption a year from now is valued at 64 percent of consumption today.

Are household savings constrained?

The relatively large discount rate implied by the drop in consumption over time seems, at first blush, inconsistent with the fact that many households chose to invest in durable goods and productive assets, such as roofs and livestock. Of course, it would be consistent if the returns to investing in metal roofs or livestock were larger than the discount rate. We therefore next quantify the returns to these investments. To this end, we conducted a separate survey of one respondent from each of 20 villages to obtain estimates for the costs of purchasing and maintaining metal and thatch roofs. The purchase of a metal roof represents an expenditure of on average USD 564, or 75 percent of the average transfer value. In addition to a store of value (roofs can be resold), a metal roof provides an investment return to households by obviating the need to replace and repair their thatched roofs, which costs on average USD 107 per year. Together, these figures imply a simple return on the investment in the roof of 23 percent (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). Similarly, treated households spend USD 85 on livestock assets, and these purchases increase their flow profit from animal husbandry by USD 1.68 per month. The implied monthly return of this investment is therefore 2 percent, and the yearly return 27 percent. Thus, many households chose to invest in asset that have yearly return that are about half of their discount rate: there is in fact a discrepancy between the discount rate implied by the drop in non-durable consumption and the savings and investment behavior.

One possible explanation for this difference is that the large effective discount rate households may find it difficult to save. The comparison between households that received monthly vs. lump-sum transfers is consistent with this interpretation: if households are both credit- and savings-constrained, 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. Conversely, recipients of a lump sum may be keen to invest it immediately into a large durable if they are not sure they can pace their non-durable consumption and save. Table 5 compares the savings and investment behavior of households that received monthly transfers to that of households that received lump-sum transfers. Indeed, we find that the value of non-land assets accumulated by monthly recipient households is significantly lower than for lump-sum recipients. In particular, monthly recipients are much less likely to acquire a metal roof compared to those who received lump-sum transfers. Instead, these households use more of the transfer for current consumption, evident in a significantly higher food security index in this group.

Thus, we find that monthly transfer recipients are indeed significantly less likely to invest durables such as metal roofs than lump-sum transfer recipients, suggesting that households may be both credit- and savings-constrained. However, 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 on their *M-Pesa* account, and even add additional funds), suggests that the savings constraint at work is more social or behavioral than purely due to the lack of access to a savings technology. The results on cortisol levels (Table 7) provide a tantalizing complement to this interpretation: we find that cortisol levels are significantly higher for households who receive the monthly transfers than for those who receive lump-sum transfers, and in fact, significantly elevated even relative to control (cf. Online Appendix). This is despite the fact that immediate pressures on the lives of these recipients have reduced, 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, while the transfers of lump-sum recipients are safely put away in metal roofs.

6 Conclusion

This paper reports the main results from an impact evaluation of unconditional cash transfers in a sample of poor households in Western Kenya. The study was designed to shed light not only on the impacts of unconditional cash transfers – a question of independent policy interest – but also on several long-standing questions in economics. In particular, by making one-time, unanticipated, large cash transfers, the intervention we study provides an ideal setting to address one of the core questions of consumer theory: how do households respond to income changes? In addition, our design enables us to ask several other questions that are central in different literatures: What is the shape of the Engel curves for different expenditure heads? Are households unitary? Are they savings- and credit-constrained? Are there externalities of transfers? The current paper only scratches the surface of the wealth of available data (which will be made publicly available).

We find that treatment households increased both consumption and savings (in the form of durable good purchases and investment in their self-employment activities). They increased food expenditures close to proportionally to overall non-durable expenditure (we estimate a food expenditure elasticity of 0.83, similar to the cross-sectional elasticity in our sample and in others), and health and education expenditures more than proportionally. Alcohol and tobacco expenditures did not increase. We find no evidence for an increase in tension within households, no significant spillover effects on non-recipient households, and no general equilibrium effects at the village level, with the single exception that we observe an increase in female empowerment at the village level. Together, these findings suggest that simple cash transfers may not have the perverse effects that some policymakers feel they would have, which has led for a clear policy preference for conditional cash transfers or in-kind transfers.

The paper contains a number of additional intriguing results, which we hope will be the object of future research. First, a puzzle is that observed returns on households' savings and investments implies a discount rate which is much higher than the the drop in consumption gains for households observed at longer intervals after the end of transfers. This discrepancy may be due to savings constraints; in line with this view, asset accumulation is significantly greater for households that received lump-sum transfers than for monthly transfer recipients. Second, contrary to previous literature and our expectation, we find no significant differences between transfers to men and transfers to women in expenditure decisions or any other outcomes. One possible account for this finding is that transfers did not affect the bargaining power of the different household members because transfers were explicitly temporary.

Finally, the treatment effects on levels of the stress hormone cortisol raise a number of intriguing questions for future research. The finding that cortisol levels were reduced to a greater extent for large compared to small transfers is significant because it constitutes the first demonstration that increases in wealth causally affect cortisol levels; however, the effect is otherwise intuitive and expected. In contrast, 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 vs. female recipients on other outcomes. This result therefore raises the question whether cortisol may track particular aspects of welfare with greater sensitivity than traditional measures. In particular, we note that a) in the cross-section, psychological well-being correlates strongly with female empowerment, and b) the treatment effect on female empowerment was larger when the female received the transfer, but not significantly so. Together, these results suggest that the difference in cortisol levels between female and male recipient households may reflect differences in female empowerment that are not visible in self-report measures such as our female empowerment index. Finally, we observe significantly lower cortisol levels when transfers are lump-sum rather than monthly. We speculate that this finding may reflect the fact that monthly transfer recipient households find it difficult to save their transfers, as evident in the fact that they do not invest in metal roofs as lump-sum recipient households do. This inability to save may result in stress and increased levels of cortisol.

These findings suggest that cortisol is a useful outcome variable for impact evaluation: the present study establishes that cortisol levels can respond to interventions, in fact sometimes to a greater degree than traditional economic outcome variables or responses to psychological questionnaires. Together with its other advantages (objectivity, correlation with psychological well-being, implications for long-term health, and ease of collection), we hope that these results will contribute to making cortisol a useful complement to existing outcome measures in impact evaluation.

References

Aggarwal, Bharat B., Shishir Shishodia, Santosh K. Sandur, Manoj K. Pandey, and Gautam Sethi. 2006. "Inflammation and cancer: how hot is the link?" *Biochemical Pharmacology* 72 (11): 1605–1621.

Aker, Jenny C., Rachid Boumnijel, Amanda McClelland, and Niall Tierney. 2011. "Zap it to me: the short-term impacts of a mobile cash transfer program." *Center for Global Development Working Paper*, no. 268.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.

Angelucci, Manuela, and Giacomo De Giorgi. 2009. "Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption?" *The American Economic Review*, pp. 486–508.

Arnetz, B B, S O Brenner, L Levi, R Hjelm, I L Petterson, J Wasserman, B Petrini, P Eneroth, A Kallner, and R Kvetnansky. 1991. "Neuroendocrine and immunologic effects of unemployment and job insecurity." *Psychotherapy and Psychosomatics* 55 (2-4): 76–80.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines." *The Quarterly Journal of Economics* 121 (2): 635–672.

Baird, Sarah, Jacobus De Hoop, and Berk Özler. 2013. "Income shocks and adolescent mental health." *Journal of Human Resources* 48 (2): 370–403.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2013. "Can basic entrepreneurship transform the economic lives of the poor?" STICERD - economic organisation and public policy discussion papers series 43, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Banerjee, A. V., E. Duflo, R. Glennerster, and D. Kothari. 2010. "Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives." *BMJ* 340:1–9.

Banerjee, Abhijit, and Esther Duflo. 2005. "Do firms want to borrow more? Testing credit constraints using a directed lending program." *BREAD Working Paper*, no. 005.

———. 2011. "Targeting the hard-core poor: An impact assessment." *MIT mimeo.*

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. 2013. "Turning a shove into a nudge? A "labeled cash transfer" for education." *NBER Working Paper*, no. 19227.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2013. "Generating skilled employment in developing countries: Experimental evidence from Uganda." *Unpublished working paper.*

Bodkin, Ronald. 1959. "Windfall income and consumption." *American Economic Review* 49 (4): 602–614.

Browning, Martin, and PierreAndre Chiappori. 1998. "Efficient intra-household allocations: A general characterization and empirical tests." *Econometrica*, pp. 1241–1278.

Browning, Martin, and Thomas F. Crossley. 2001. "The life-cycle model of consumption and saving." *Journal of Economic Perspectives* 15 (3): 3–22.

Brune, Lasse, Xavier Gine, Jessica Goldberg, and Dean Yang. 2011. "Commitments to save: A field experiment in rural Malawi." Policy research working paper series 5748, The World Bank.

Burbidge, John B., Lonnie Magee, and A. Leslie Robb. 1988. "Alternative transformations to handle extreme values of the dependent variable." *Journal of the American Statistical Association* 83 (401): 123–127.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust inference with multiway clustering." *Journal of Business & Economic Statistics* 29, no. 2.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping institutions: Evidence on aid impacts using a preanalysis plan." *The Quarterly Journal of Economics* 127 (4): 1755–1812.

Chemin, Matthieu, De Laat, Joost, and Johannes Haushofer. 2013. "Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in Kenya." SSRN scholarly paper ID 2294171, Social Science Research Network, Rochester, NY.

Cohen, Sheldon, William J Doyle, and Andrew Baum. 2006. "Socioeconomic status is associated with stress hormones." *Psychosomatic medicine* 68 (3): 414–420.

Cohen, Sheldon, Joseph E Schwartz, Elissa Epel, Clemens Kirschbaum, Steve Sidney, and Teresa Seeman. 2006. "Socioeconomic status, race, and diurnal cortisol decline in the Coronary Artery Risk Development in Young Adults (CARDIA) Study." *Psychosomatic Medicine* 68 (1): 41–50.

Cornelisse, Sandra, Van Ast, Vanessa, Johannes Haushofer, Maayke Seinstra, and Marian Joels. 2013. "Time-dependent effect of hydrocortisone administration on intertemporal choice." SSRN scholarly paper ID 2294189, Social Science Research Network, Rochester, NY.

Coussens, Lisa M., and Zena Werb. 2002. "Inflammation and cancer." *Nature* 420 (6917): 860–867.

Cunha, Jesse M., Giacomo De Giorgi, and Seema Jayachandran. 2011. "The price effects of cash versus in-kind transfers." Technical Report, National Bureau of Economic Research.

Dasgupta, Partha, and Debraj Ray. 1986. "Inequality as a determinant of malnutrition and unemployment: Theory." *Economic Journal* 96 (384): 1011–1034.

Deaton, Angus. 1992. *Understanding consumption.* Oxford University Press.

———. 2002. "Poverty and inequality in India: A reexamination." *Economic and Political Weekly*, pp. 3729–3748.

Deaton, Angus, and Shankar Subramanian. 1996. "The demand for food and calories." *Journal of Political Economy*, pp. 133–162.

Deaton, Angus, and Alessandro Tarozzi. 2005. Chapter Prices and poverty in India of *The Great Indian Poverty Debate*, edited by Angus Deaton and Valerie Kozel. Macmillan.

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to capital in microenterprises: evidence from a field experiment." *The Quarterly Journal of Economics* 123 (4): 1329–1372.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Pariente, and Vincent Pons. 2011. "Happiness on tap: Piped water adoption in urban Morocco." Working paper 16933, National Bureau of Economic Research.

Doss, Cheryl R. 1996. "Testing among models of intrahousehold resource allocation." *World Development* 24 (10): 1597–1609.

Duflo, Esther. 2003. "Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa." *The World Bank Economic Review* 17 (1): 1–25.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2008. "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *American Economic Review* 98 (2): 482–488.

Duflo, Esther, and Emmanuel Saez. 2003. "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment." *The Quarterly Journal of Economics* 118 (3): 815–842.

Duflo, Esther, and Christopher Udry. 2004. "Intrahousehold resource allocation in Cote D'Ivoire: Social norms, separate accounts and consumption choices." *NBER Working Paper*, no. 10498.

Dupas, Pascaline, and Jonathan Robinson. 2013a. "Savings constraints and microenterprise: Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics* 5 (1): 163–192.

———. 2013b. "Why don't the poor save more? Evidence from health savings experiments." *American Economic Review* 103 (4): 1138–1171.

Dynarski, Susan, Jonathan Gruber, Robert Moffitt, and Gary Burtless. 1997. "Can families smooth variable earnings." *Brookings Papers on Economic Activity*, no. 1:229–303.

Efron, Bradley, and Robert Tibshirani. 1993. *An introduction to the bootstrap*. CRC press.

Fafchamps, Marcel, David McKenzie, Simon R. Quinn, and Christopher Woodruff. 2011. "When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana." Technical Report, National Bureau of Economic Research.

Fernald, Lia, and Megan R Gunnar. 2009. "Effects of a poverty-alleviation intervention on salivary cortisol in very low-income children." *Social Science & Medicine (1982)* 68 (12): 2180–2189.

Ferracuti, S, S Seri, D Mattia, and G Cruccu. 1994. "Quantitative EEG modifications during the Cold Water Pressor Test: hemispheric and hand differences." *International Journal of Psychophysiology* 17 (3): 261–268.

Friedman, Milton. 1957. *A theory of the consumption.* Princeton University Press.

Goldstein, Jacib. 2013. "Is it nuts to give to the poor without strings attached?" *New York Times.*

Hall, Robert E., and Frederic S. Mishkin. 1982. "The sensitivity of consumption to transitory income: Estimates from panel data on households." *Econometrica* 50 (2): 461–481.

Hammen, Constance. 2005. "Stress and depression." *Annual Review of Clinical Psychology* 1:293–319.

Haushofer, Johannes. 2011. "Neurobiological poverty traps." *Working Paper.*

Haushofer, Johannes, Sandra Cornelisse, Marian Joels, Tobias Kalenscher, and Ernst Fehr. 2011. "Low income is associated with high baseline levels and low stress reactivity of cortisol, but not alpha amylase." *Working Paper.*

Haushofer, Johannes, and Ernst Fehr. 2013. "The psychology and neurobiology of poverty." *Submitted.*

Haushofer, Johannes, Ernst Fehr, and Daniel Schunk. 2013. "Negative income shocks increase discount rates." *Working Paper.*

Heckman, James. 1992. In *Randomization and social program*, edited by C. Manski and S. Garfinkle. Cambridge, MA: Harvard University Press.

Hoddinott, John, and Lawrence Haddad. 1995. "Does female income share influence household expenditures? Evidence from Côte D'ivoire." *Oxford Bulletin of Economics and Statistics* 57 (1): 77–96.

Holsboer, F. 2000. "The corticosteroid receptor hypothesis of depression." *Neuropsychopharmacology* 23 (5): 477–501.

Horton, R., and R. Smith. 1999. "Time to register for randomized trials." *British Medical Journal* 319:865.

Jack, William, and Tavneet Suri. 2013. "Risk sharing and transaction costs: Evidence from Kenya's mobile money revolution." *Working Paper.*

Jappelli, Tullio, and Luigi Pistaferri. 2010. "The consumption response to income changes." *Annual Review of Economics* 2 (1): 479–506.

Jensen, Robert T., and Nolan H. Miller. 2008a. "Giffen behavior and subsistence consumption." *American Economic Review* 98 (4): 1553.

———. 2008b. "The impact of food prices increases on caloric intake in China." *Agricultural Economics* 39 (1): 465–476.

Kahneman, Daniel, and Angus Deaton. 2010. "High income improves evaluation of life but not emotional well-being." *Proceedings of the National Academy of Sciences* 107 (38): 16489–16493.

Karlan, Dean, and Jonathan Zinman. 2011. "Microcredit in theory and practice: Using randomized credit scoring for impact evaluation." *Science* 332 (6035): 1278–1284.

Kiecolt-Glaser, Janice K., Kristopher J. Preacher, Robert C. MacCallum, Cathie Atkinson, William B. Malarkey, and Ronald Glaser. 2003. "Chronic stress and age-related increases in the proinflammatory cytokine IL-6." *Proceedings of the National Academy of Sciences* 100 (15): 9090–9095.

Kirschbaum, C., C. J. Strasburger, and J. Langkrär. 1993. "Attenuated cortisol response to psychological stress but not to CRH or ergometry in young habitual smokers." *Pharmacology Biochemistry and Behavior* 44 (3): 527–531.

Kirschbaum, Clemens, and Dirk H. Hellhammer. 1989. "Salivary cortisol in psychobiological research: an overview." *Neuropsychobiology* 22 (3): 150–169.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75 (1): 83–119.

Knorr, Ulla, Maj Vinberg, Lars V. Kessing, and Jørn Weeterslev. 2010. "Salivary cortisol in depressed patients versus control persons: A systematic review and meta-analysis." *Psychoneuroendocrinology* 35 (9): 1275–1286.

Kochar, Anjini. 1995. "Explaining household vulnerability to idisyncratic income shocks." *American Economic Review* 85 (2): 159–164.

Krueger, Dirk, and Fabrizio Perri. 2005. "Understanding consumption smoothing: Evidence from the US consumer expenditure data." *Journal of the European Economic Association* 3 (2–3): 340 349.

———. 2006. "Does income inequality lead to consumption inequality? Evidence and theory." *The Review of Economic Studies* 73 (1): 163–193.

———. 2010. "How do households respond to income shocks."

Lee, Soohyung, and Azeem M. Shaikh. 2013. "Multiple testing and heterogeneous treatment effects: Re-evaluating the effect of Progresa on school enrollment." *Journal of Applied Econometrics.*

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales. 1997. "Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit." *Journal of Human Resources* 32, no. 3.

MacKinnon, James G., and Lonnie Magee. 1990. "Transforming the dependent variable in regression models." *International Economic Review* 31(2):315–39.

Mani, Anand, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty impedes cognitive function." *Science* 341 (6149): 976–980.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99 (2): 210–221.

Modigliani, Franco, and Richard Brumberg. 1954. Chapter Utility analysis and the consumption function: An attempt at integration of *Post Keynesian Economics*, edited by Kenneth Kurihara. Rutgers University Press.

Paxson, Christina H. 1993. "Consumption and income seasonality in Thailand." *Journal of Political Economy*, pp. 39–72.

Pence, Karen M. 2006. "The role of wealth transformations: an application to estimating the effect of tax incentives on saving." *Contributions to Economic Analysis and Policy*, vol. 5(1).

Pepper, John V. 2002. "Robust inferences from random clustered samples: an application using data from the panel study of income dynamics." *Economics Letters* 75 (3): 341–345.

Romano, Joseph P., and Michael Wolf. 2005. "Exact and approximate stepdown methods for multiple hypothesis testing." *Journal of the American Statistical Association* 100 (469): 94–108.

Rosenthal, Robert. 1979. "The file drawer problem and tolerance for null results." *Psychological Bulletin* 86 (3): 638–641.

Rosenzweig, Mark R., and Kenneth Wolpin. 1982. "Governmental interventions and household behavior ina developing country." *Journal of Development Economics* 10:209–225.

Ross, Russell. 1999. "Atherosclerosis — An Inflammatory Disease." *New England Journal of Medicine* 340 (2): 115–126.

Sacks, Daniel W., Betsey Stevenson, and Justin Wolfers. 2012. "The new stylized facts about income and subjective well-being." *Emotion* 12 (6): 1181.

Shah, Anuj K., Sendhil Mullainathan, and Eldar Shafir. 2012. "Some consequences of having too little." *Science* 338 (6107): 682–685.

Simes, R.J. 1986. "Publication bias: The case for and international registry of clinical trials." *Journal of Clinical Oncology* 4:1529–1541.

Steptoe, A., P. J. Feldman, S. Kunz, N. Owen, G. Willemsen, and M. Marmot. 2002. "Stress responsivity and socioeconomic status. A mechanism for increased cardiovascular disease risk?" *European Heart Journal* 23 (22): 1757–1763.

Steptoe, Andrew, Gonneke Willemsen, Natalie Owen, Louise Flower, and Vidya Mohamed-Ali. 2001. "Acute mental stress elicits delayed increases in circulating inflammatory cytokine levels." *Clinical Science* 101 (2): 185–192.

Stevenson, Betsey, and Justin Wolfers. 2008. "Economic growth and subjective well-being: Reassessing the Easterlin paradox." Technical Report, National Bureau of Economic Research.

Straub, Rainer H. 2006. "Bottom-up and top-down signaling of IL-6 with and without habituation?" *Brain, Behavior, and Immunity* 20 (1): 37–39.

Strauss, John, and Duncan Thomas. 1995. Chapter Human resources: Empirical modeling of household and family decisions of *Handbook of Development Economics*, edited by Jere Behrman and T.N. Srinivasan. Elsevier.

Team, The Kenya CT-OVC Evaluation. 2012a. "The impact of Kenya's Cash Transfer for Orphans and Vulnerable Children on human capital." *Journal of Development Effectiveness* 4 (1): 38–49.

———. 2012b. "The impact of the Kenya Cash Transfer Program for Orphans and Vulnerable Children on household spending." *Journal of Development Effectiveness* 4 (1): 9–37.

Thomas, D. 1989. "Intra-household resource allocation: An inferential approach." Paper 586, Yale - Economic Growth Center.

Townsend, Richard. 1994. "Risk and insurance in village India." *Econometrica* 62 (3): 539–591.

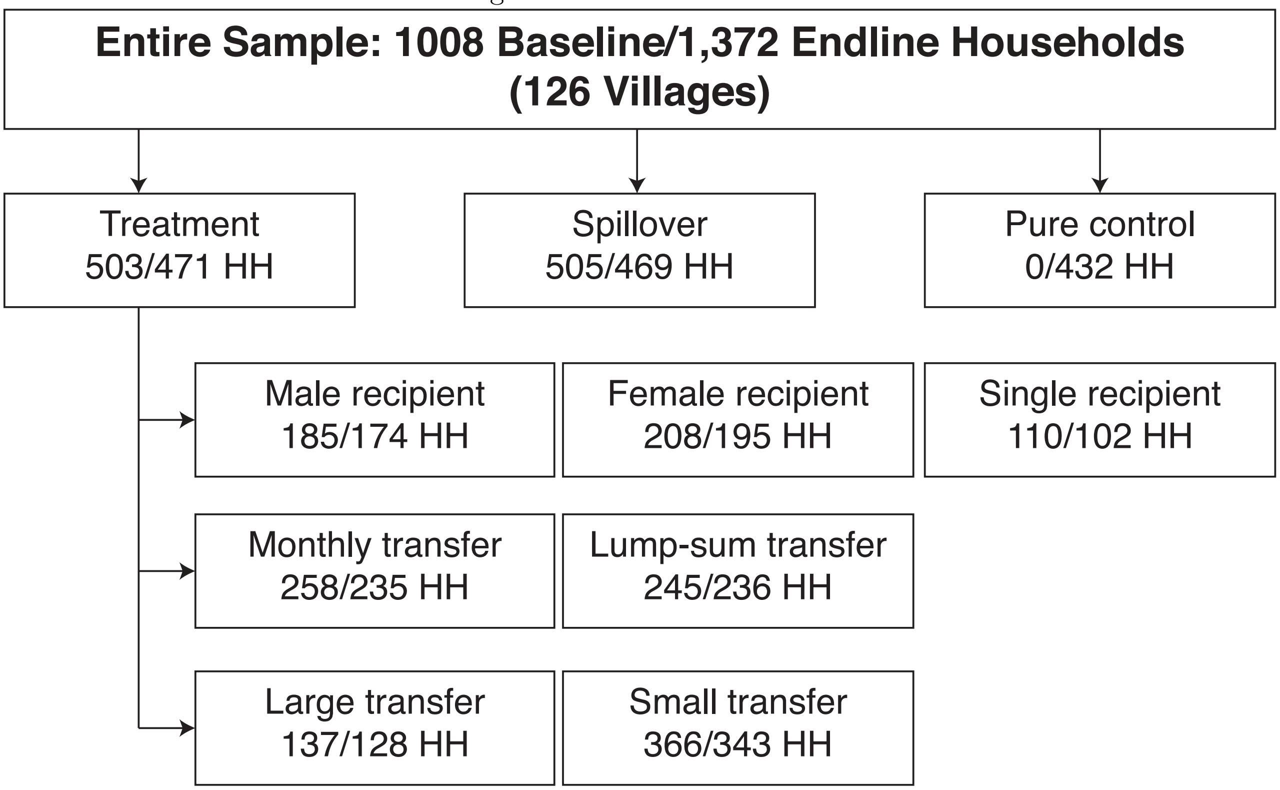
Udry, Christopher. 1996. "Gender, agricultural production, and the theory of the household." *Journal of Political Economy*, pp. 1010–1046.

Westfall, P. H., S. S. Young, and S. Paul Wright. 1993. "On adjusting P-values for multiplicity." *Biometrics* 49 (3): 941–945.

Wilckens, T. 1995. "Glucocorticoids and immune function: physiological relevance and pathogenic potential of hormonal dysfunction." *Trends in Pharmacological Sciences* 16 (6): 193–197 (June).

Zwane, Alix Peterson, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, Florencia Devoto, and Crep. 2011. "Being surveyed can change later behavior and related parameter estimates." *Proceedings of the National Academy of Sciences* 108 (5): 1821–1826.

Figure 1: Treatment arms



Notes: Diagram of treatment arms. Numbers designate baseline/endline number of households in each treatment arm. Pure control households were added at endline to allow identification of spillover effects. Male and female recipient was randomized only for households with co-habitating couples. Large transfers were administered by making additional transfers to households that had previously been assigned to treatment.

Figure 2: Timeline of study

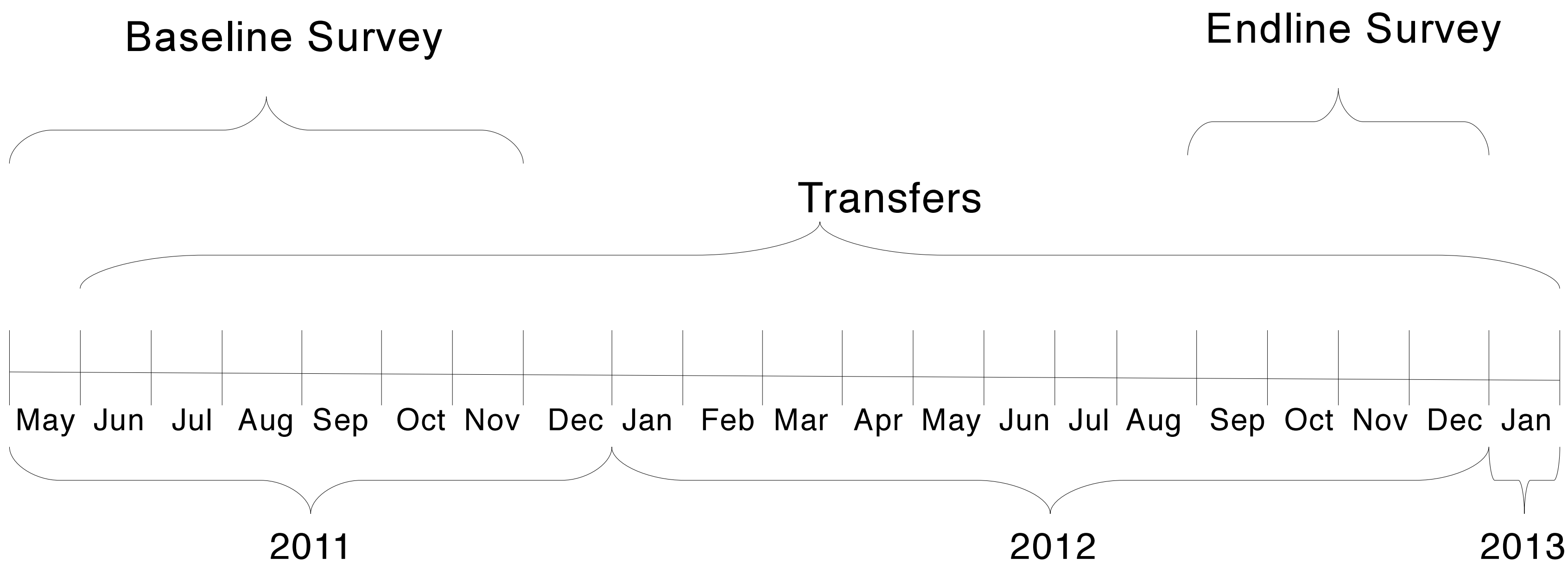
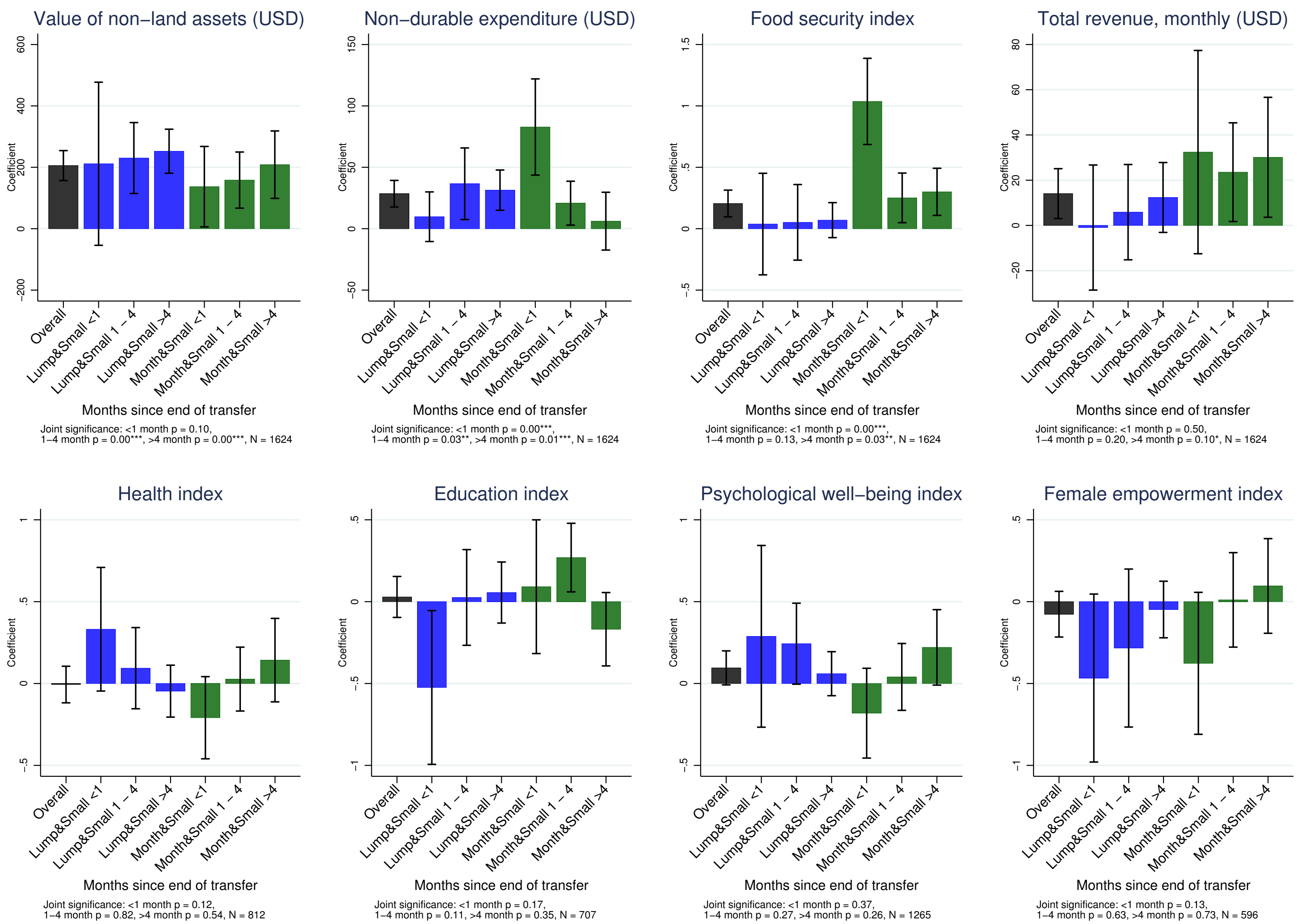
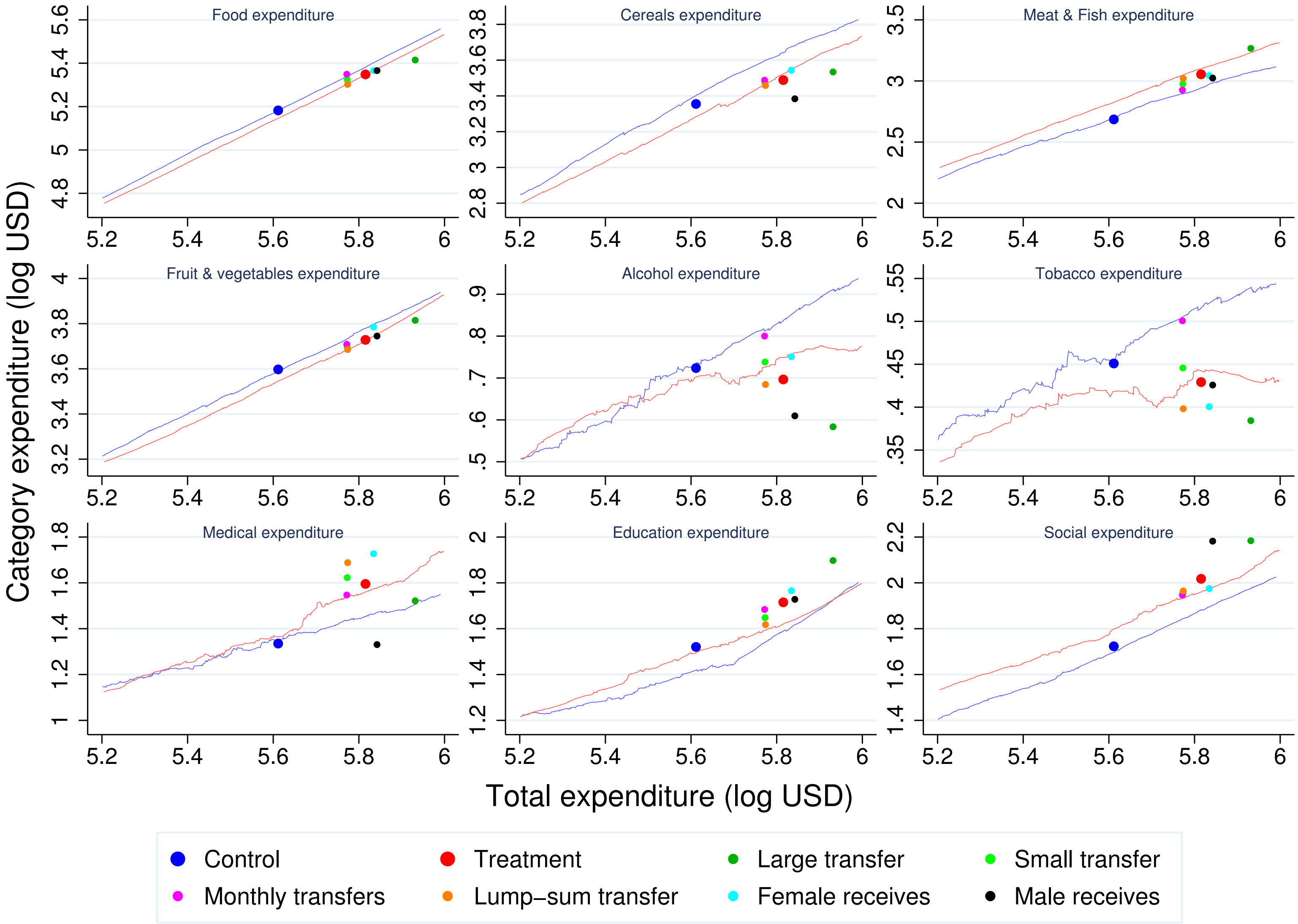


Figure 3: Treatment effects on index variables over time



Notes: Treatment effects on index variables over time. Shown are coefficient estimates and error bars representing 95 pct. confidence intervals, separately for the treatment effect of transfers ending less than 1 month ago, 1–4 months ago, and more than 4 months ago. The results are shown separately for the monthly and lump-sum transfer groups; Wald tests of joint significance across these groups are presented below each panel. Assets, consumption, and income are coded in USD (PPP); the other variables are indices in z-score units, with higher values corresponding to "positive" outcomes.

Figure 4: Cross-sectional and experimental Engel curves for different categories of consumption and expenditure



Notes: Cross-sectional and experimental Engel curves for different categories of consumption and expenditure. We plot the log of total monthly nondurable expenditure on the horizontal axis, and the log of monthly expenditure on sub-categories on the vertical axis. Because both axes are on log scales, slopes correspond to elasticities. Total expenditure includes spending on food, education, health care, and social expenses, but not investment in business and agriculture and spending on durables. The red and blue lines represent cross-sectional Engel curves, estimated with local linear regressions of the log of each category's expenditure on the log of total expenditure for the treatment and control groups at endline, respectively. The large red and blue dots represent the experimental Engel curve; the blue dot shows the average total and category expenditure for the control group, and the red dot for the control group. The smaller colored dots represent subgroups of the treatment group (transfers to male vs. female, monthly vs. lump-sum transfers, and large vs. small transfers).

Table 1: Treatment effects: Index variables

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Value of non-land assets (USD)	477.66	278.52***	−18.73	−66.19	−74.97*	252.84***	1372
	(389.23)	(25.44)	(21.07)	(47.25)	(42.92)	(45.94)	
		[0.00]***	[0.90]	[0.72]	[0.44]	[0.00]***	
Non-durable expenditure (USD)	157.40	36.18***	−7.53	−2.74	−4.40	20.37*	1372
	(82.18)	(5.91)	(7.24)	(10.35)	(10.82)	(10.55)	
		[0.00]***	[0.90]	[0.99]	[1.00]	[0.29]	
Total revenue, monthly (USD)	48.98	16.64***	−5.23	5.30	16.20	−1.64	1372
	(90.52)	(5.93)	(5.67)	(10.61)	(11.11)	(8.96)	
		[0.03]**	[0.90]	[0.99]	[0.57]	[0.84]	
Food security index	−0.00	0.25***	0.04	0.05	0.26**	0.16	1372
	(1.00)	(0.06)	(0.10)	(0.09)	(0.11)	(0.10)	
		[0.00]***	[0.90]	[0.99]	[0.12]	[0.30]	
Health index	−0.00	−0.04	−0.08	0.09	0.01	−0.07	1372
	(1.00)	(0.06)	(0.08)	(0.09)	(0.10)	(0.09)	
		[0.78]	[0.90]	[0.86]	[1.00]	[0.78]	
Education index	0.00	0.08	−0.00	0.05	−0.05	0.04	1174
	(1.00)	(0.06)	(0.08)	(0.09)	(0.10)	(0.09)	
		[0.50]	[0.97]	[0.99]	[0.98]	[0.83]	
Psychological well-being index	−0.00	0.20***	0.08	0.16*	−0.10	0.35***	2140
	(1.00)	(0.06)	(0.07)	(0.09)	(0.10)	(0.10)	
		[0.03]**	[0.18]	[0.99]	[1.00]	[0.00]***	
Female empowerment index	−0.00	−0.01	0.23**	0.16	0.04	0.21**	1010
	(1.00)	(0.07)	(0.09)	(0.10)	(0.12)	(0.11)	
		[0.93]	[0.05]*	[0.67]	[1.00]	[0.29]	
Joint test (p-value)		0.00***	0.13	0.41	0.01**	0.00***	

Notes: OLS estimates of treatment effects. Outcome variables are listed on the left. Assets, consumption, and income are coded in KES; the other variables are indices in z-score units, with higher values corresponding to "positive" outcomes. For each outcome variable, we report the coefficients of interest and their standard errors in parentheses. FWER-corrected standard errors 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) 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 co-habitating couples for the female empowerment index, and households with school-age children for the education index. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 2: Treatment effects: Consumption

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Food total (USD)	104.46	19.60***	−3.48	−2.26	1.76	7.71	1372
	(58.50)	(4.22)	(4.66)	(7.43)	(7.51)	(7.62)	
Food own production (USD)	13.64	2.45**	−2.09*	0.16	3.94**	−0.23	1372
	(14.79)	(0.96)	(1.18)	(1.72)	(1.77)	(1.48)	
Food bought (USD)	90.82	16.98***	−1.39	−3.11	−3.03	7.49	1372
	(52.77)	(3.81)	(4.31)	(6.61)	(6.73)	(6.84)	
Cereals (USD)	22.55	2.24**	0.30	0.24	−1.24	2.45	1372
	(17.18)	(1.14)	(1.58)	(1.87)	(1.87)	(2.08)	
Meat & fish (USD)	12.97	5.10***	−0.35	0.76	−3.12	2.41	1372
	(13.75)	(1.02)	(1.22)	(1.83)	(1.95)	(1.64)	
Fruit & vegetables (USD)	23.50	3.46***	0.20	−0.95	0.13	2.29	1372
	(17.06)	(1.15)	(1.39)	(1.96)	(2.05)	(1.99)	
Dairy (USD)	7.26	1.71***	−0.16	−0.73	0.82	0.49	1372
	(9.43)	(0.64)	(0.74)	(1.10)	(1.09)	(1.09)	
Fats (USD)	6.84	0.80**	0.01	−0.28	−0.27	0.91	1372
	(5.51)	(0.37)	(0.46)	(0.62)	(0.64)	(0.58)	
Sugars (USD)	11.25	1.05**	−0.52	−0.53	0.10	0.41	1372
	(7.18)	(0.48)	(0.56)	(0.81)	(0.84)	(0.78)	
Other food (USD)	42.42	5.98***	−0.36	−1.55	−0.86	3.31	1372
	(28.28)	(1.94)	(2.40)	(3.24)	(3.23)	(3.43)	
Alcohol (USD)	6.38	−0.93	−0.41	1.50	1.00	−1.55	1372
	(16.56)	(1.00)	(1.26)	(1.64)	(1.65)	(1.35)	
Tobacco (USD)	1.52	−0.16	−0.00	0.11	0.43	−0.31	1372
	(4.13)	(0.22)	(0.29)	(0.34)	(0.34)	(0.30)	
Medical expenditure past month (USD)	6.56	2.83***	1.52	2.06	−1.49	−0.35	1372
	(13.17)	(0.98)	(0.93)	(1.86)	(1.87)	(1.73)	
Medical expenditure, children (USD)	3.52	0.66	1.03*	0.63	−0.37	−0.10	1203
	(8.52)	(0.60)	(0.60)	(1.06)	(1.09)	(0.97)	
Education expenditure (USD)	4.71	1.08**	0.32	0.44	−0.10	1.10	1372
	(8.68)	(0.51)	(0.61)	(0.89)	(0.88)	(0.92)	
Social expenditure (USD)	4.36	2.46***	−1.42***	−2.06**	−0.46	0.67	1372
	(5.38)	(0.49)	(0.46)	(0.98)	(1.01)	(0.90)	
Other expenditure (USD)	34.36	10.06***	−3.72	−2.05	−3.56	11.76***	1372
	(24.62)	(1.74)	(2.27)	(3.05)	(3.17)	(3.01)	
Non-durable expenditure (USD)	157.40	36.18***	−7.53	−2.74	−4.40	20.37*	1372
	(82.18)	(5.91)	(7.24)	(10.35)	(10.82)	(10.55)	
Joint test (p-value)		0.00***	0.15	0.81	0.39	0.03**	

Notes: OLS estimates of treatment and spillover effects. Outcome variables are listed on the left. All variables are reported in PPP adjusted USD. Food bought includes all sub-categories except alcohol and tobacco. Education expenditures include tuition and other costs (e.g., uniform, supplies). Social expenditures include charitable donations, dowry, fees paid to village elder or chiefs, religious ceremonies, weddings, funerals and recreation (e.g., books, music). Other expenditures include airtime, travel and transportation, clothing, personal items (e.g., toiletries), household items (e.g., soap, candles), firewood, electricity and water. Non-durable expenditures are the sum of expenditures on food (own production and purchased), alcohol and tobacco, medical, education, social activities and other goods. Variables in PPP adjusted USD are topcoded for the highest 1 percent 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) that of large compared to small transfers. The unit of observation is the household. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 3: Elasticity of different expenditure categories with respect to total expenditure

	Entire sample			Monthly vs. lump-sum transfers			Female vs. male recipient			Large vs. small transfer		
	(1) OLS	(2) IV	(3) Hausman p-value	(4) Monthly transfers	(5) Lump-sum transfers	(6) Difference p-value	(7) Female recipient	(8) Male recipient	(9) Difference p-value	(10) Large transfers	(11) Small transfers	(12) Difference p-value
Food total	1.00***	0.83***	0.05**	1.19***	0.69***	0.08*	0.88***	0.77***	0.62	0.68***	1.04***	0.10
	(0.02)	(0.08)		(0.20)	(0.20)		(0.14)	(0.12)		(0.12)	(0.16)	
Food own production (USD)	0.92***	1.10***	0.53	2.64**	0.38	0.10*	1.11**	1.29***	0.81	0.63*	1.72**	0.21
	(0.09)	(0.31)		(1.17)	(0.64)		(0.54)	(0.43)		(0.35)	(0.74)	
Food bought (USD)	1.03***	0.87***	0.18	1.03***	0.84***	0.57	0.90***	0.76***	0.58	0.76***	1.03***	0.30
	(0.04)	(0.10)		(0.22)	(0.21)		(0.17)	(0.14)		(0.12)	(0.21)	
Cereals (USD)	1.20***	0.75**	0.29	0.98	0.63	0.77	1.43***	−0.25	0.05*	0.64*	0.89	0.77
	(0.09)	(0.33)		(0.83)	(0.68)		(0.54)	(0.62)		(0.35)	(0.70)	
Meat & fish (USD)	1.17***	2.07***	0.01**	1.61*	2.52***	0.48	2.04***	1.35***	0.40	2.02***	2.14***	0.90
	(0.09)	(0.37)		(0.89)	(0.77)		(0.58)	(0.44)		(0.39)	(0.79)	
Fruit & vegetables (USD)	0.95***	0.76***	0.30	1.11**	0.49	0.35	1.10***	0.58*	0.29	0.74***	0.78*	0.93
	(0.06)	(0.19)		(0.45)	(0.43)		(0.30)	(0.31)		(0.22)	(0.40)	
Dairy (USD)	1.44***	1.41***	0.95	3.03**	0.97	0.21	1.65**	1.71***	0.95	0.45	2.70**	0.09*
	(0.11)	(0.45)		(1.28)	(0.88)		(0.77)	(0.63)		(0.59)	(1.09)	
Fats (USD)	0.89***	0.62***	0.32	0.49	0.61	0.89	0.79**	0.68**	0.85	0.76***	0.43	0.59
	(0.07)	(0.24)		(0.56)	(0.48)		(0.38)	(0.33)		(0.25)	(0.51)	
Sugars (USD)	0.89***	0.68***	0.46	1.14**	0.55	0.45	0.60	0.75**	0.79	0.43	1.00*	0.39
	(0.08)	(0.25)		(0.56)	(0.47)		(0.38)	(0.34)		(0.30)	(0.52)	
Other food (USD)	1.14***	0.80***	0.16	1.06***	0.66*	0.52	0.98***	0.57**	0.32	0.69***	0.95***	0.54
	(0.06)	(0.18)		(0.40)	(0.38)		(0.28)	(0.26)		(0.18)	(0.36)	
Alcohol (USD)	0.53***	−0.13	0.36	0.98	−0.59	0.43	0.27	−0.61	0.54	−0.57	0.47	0.46
	(0.13)	(0.56)		(1.46)	(1.14)		(0.98)	(0.83)		(0.59)	(1.16)	
Tobacco (USD)	0.24**	−0.19	0.35	0.07	−0.10	0.89	−0.61	−0.19	0.66	−0.49	0.19	0.46
	(0.09)	(0.36)		(0.85)	(0.70)		(0.66)	(0.53)		(0.41)	(0.74)	
Medical expenditure past month (USD)	0.36***	1.47**	0.12	1.27	2.61*	0.54	3.16**	−0.99	0.01**	0.35	2.96*	0.15
	(0.13)	(0.58)		(1.42)	(1.37)		(1.29)	(0.97)		(0.67)	(1.57)	
Medical expenditure, children (USD)	0.18	0.86*	0.38	0.38	1.86	0.41	1.93**	−0.95	0.05*	0.17	1.89	0.25
	(0.12)	(0.47)		(1.12)	(1.14)		(0.97)	(0.95)		(0.54)	(1.27)	
Education expenditure (USD)	0.75***	0.84**	0.88	1.22	0.55	0.62	1.12*	0.42	0.45	0.88**	0.78	0.92
	(0.10)	(0.37)		(0.93)	(0.77)		(0.63)	(0.52)		(0.42)	(0.80)	
Social expenditure (USD)	0.74***	1.60***	0.02**	1.48*	1.61**	0.91	0.71	2.44***	0.06*	1.71***	1.45**	0.78
	(0.07)	(0.35)		(0.87)	(0.71)		(0.57)	(0.65)		(0.42)	(0.73)	

Notes: Elasticity of different expenditure and consumption categories with respect to total expenditure, for different subsets of the treatment group. Column (1) presents cross-sectional OLS estimates in the control group; column (2) presents IV estimates across both treatment and control groups when total expenditure is instrumented with treatment. Column (3) shows the p-value of the Hausman test comparing OLS and IV specifications. Columns (4) and (5) present IV estimates for the effect of monthly and lump-sum transfers, respectively; column (6) shows the p-value of the difference. Analogously for the remaining columns.

Table 4: Treatment effects: Food security

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Meals skipped (adults, # last month)	4.38	−0.99***	0.25	−0.14	−0.50	−0.30	1372
	(5.75)	(0.35)	(0.49)	(0.48)	(0.56)	(0.52)	
Whole days without food (adults, # last month)	0.87	−0.27*	0.07	0.01	−0.20	0.26	1372
	(2.73)	(0.15)	(0.17)	(0.19)	(0.15)	(0.28)	
Meals skipped (children, # last month)	2.03	−0.59**	0.61	0.23	−0.37	−0.69**	1203
	(4.48)	(0.27)	(0.42)	(0.34)	(0.48)	(0.33)	
Whole days without food (children, # last month)	0.33	−0.14*	0.13	−0.04	−0.18	−0.08	1203
	(1.37)	(0.08)	(0.10)	(0.08)	(0.11)	(0.10)	
Eat less preferred/cheaper foods (# last month)	8.17	−0.99**	0.87	0.37	−0.50	0.43	1372
	(7.69)	(0.46)	(0.73)	(0.72)	(0.78)	(0.73)	
Rely on help from others for food (# last month)	1.87	−0.08	−0.28	0.50	−0.20	0.63	1372
	(3.86)	(0.25)	(0.36)	(0.34)	(0.41)	(0.45)	
Purchase food on credit (# last month)	3.12	−0.44*	−0.43	0.08	−0.43	−0.67*	1372
	(4.57)	(0.26)	(0.44)	(0.38)	(0.42)	(0.38)	
Hunt, gather wild food, harvest prematurely (# last month)	4.10	0.04	−0.15	−0.74	−0.67	0.73	1372
	(6.78)	(0.41)	(0.70)	(0.68)	(0.67)	(0.71)	
Beg because not enough food in the house (# last month)	0.31	−0.05	−0.08	0.01	−0.05	−0.03	1372
	(0.80)	(0.05)	(0.08)	(0.07)	(0.08)	(0.07)	
All members usually eat two meals (dummy)	0.90	0.03*	0.03	0.01	0.02	0.02	1372
	(0.29)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	
All members usually eat until content (dummy)	0.79	0.04*	−0.04	−0.02	0.06	0.06	1372
	(0.41)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	
Number of times ate meat or fish (last week)	2.41	0.49***	−0.04	0.49*	0.60**	0.29	1372
	(2.07)	(0.14)	(0.22)	(0.25)	(0.27)	(0.23)	
Enough food in the house for tomorrow? (dummy)	0.36	0.07**	0.02	−0.02	0.11**	0.13**	1372
	(0.48)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	
Respondent slept hungry (last week, dummy)	0.23	−0.07***	0.03	−0.02	−0.07*	−0.08***	1372
	(0.42)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	
Respondent ate protein (last 24h, dummy)	0.29	0.07**	−0.03	0.04	0.08	0.07	1372
	(0.46)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	
Proportion of HH who ate protein (last 24h)	0.27	0.07**	−0.03	0.05	0.06	0.04	1372
	(0.42)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	
Proportion of children who ate protein (last 24h)	0.26	0.07**	−0.03	0.05	0.04	0.03	1203
	(0.42)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	
Food security index (children)	0.00	0.21***	−0.16*	0.05	0.16	0.14	1203
	(1.00)	(0.06)	(0.09)	(0.09)	(0.11)	(0.09)	
Food security index	−0.00	0.25***	0.04	0.05	0.26**	0.16	1372
	(1.00)	(0.06)	(0.10)	(0.09)	(0.11)	(0.10)	
Joint test (p-value)		0.00***	0.16	0.76	0.71	0.01**	

Notes: OLS estimates of treatment and spillover effects. Outcome variables are listed on the left. Variables indicating frequency of events in the last month are based on midpoint of ranges measured (2-4 = 3, 5-10 = 7.5, >10 = 20). Number of time ate meat or fish in last week is measured directly. The remaining outcomes are indicator variables or proportions, as indicated. 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) that of large compared to small transfers. The unit of observation is the household. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 5: Treatment effects: Assets

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Value of non-land assets (USD)	477.66	278.52***	−18.73	−66.19	−74.97*	252.84***	1372
	(389.23)	(25.44)	(21.07)	(47.25)	(42.92)	(45.94)	
Value of livestock (USD)	166.82	84.52***	−11.61	7.86	2.07	63.19**	1372
	(240.59)	(15.24)	(16.88)	(29.35)	(27.46)	(28.30)	
Value of cows (USD)	101.78	56.78***	−0.53	20.07	−13.46	43.79*	1372
	(211.82)	(13.88)	(14.89)	(26.89)	(25.40)	(25.88)	
Value of small livestock (USD)	25.30	15.15***	−4.75	−10.30	6.84	20.09***	1372
	(49.67)	(3.30)	(3.88)	(6.29)	(5.88)	(5.88)	
Value of birds (USD)	39.74	11.98***	−6.33*	−3.33	8.63*	−0.70	1372
	(40.80)	(2.77)	(3.40)	(4.92)	(5.17)	(4.50)	
Value of durable goods (USD)	207.30	53.27***	−8.74	−1.01	−8.42	63.97***	1372
	(130.60)	(8.68)	(10.88)	(14.53)	(14.34)	(15.70)	
Value of furniture (USD)	138.11	34.67***	0.33	1.48	0.65	46.08***	1372
	(89.29)	(6.06)	(7.59)	(10.26)	(10.18)	(11.40)	
Value of agricultural tools (USD)	10.77	1.61	−0.97	−2.22	−1.00	4.12**	1372
	(14.08)	(1.00)	(1.04)	(1.85)	(1.58)	(2.06)	
Value of radio/TV (USD)	9.73	2.84**	−2.12*	−0.90	2.17	0.64	1372
	(17.09)	(1.11)	(1.11)	(2.02)	(2.06)	(1.81)	
Value of bike/motorbike (USD)	21.06	2.92	−2.03	−0.48	−1.27	2.53	1372
	(35.01)	(2.27)	(2.34)	(4.21)	(3.82)	(3.82)	
Value of appliances (USD)	3.78	0.70*	−0.04	−0.04	0.22	0.52	1372
	(5.22)	(0.36)	(0.37)	(0.58)	(0.58)	(0.67)	
Value of cell phone (USD)	23.86	12.71***	−3.89*	−0.24	−2.92	7.37***	1372
	(24.85)	(1.53)	(2.00)	(2.41)	(2.54)	(2.49)	
Value of savings (USD)	10.93	10.22***	1.62	−3.43	1.81	10.22**	1372
	(29.09)	(2.48)	(2.13)	(5.10)	(4.63)	(5.07)	
Land owned (acres)	1.31	0.03	−0.08	−0.12	0.01	0.35	1372
	(1.88)	(0.14)	(0.16)	(0.18)	(0.18)	(0.31)	
Has non-thatched roof (dummy)	0.16	0.23***	0.00	−0.12**	−0.12**	0.23***	1372
	(0.37)	(0.03)	(0.00)	(0.05)	(0.05)	(0.05)	
Joint test (p-value)		0.00***	0.32	0.40	0.25	0.00***	

Notes: OLS estimates of treatment effects. Outcome variables are listed on the left. Value of asset variables are reported in PPP adjusted USD. Has non-thatched roof is an indicator variable. Value of livestock includes value of cows, small livestock (e.g., goats) and birds. Value of durable goods is the sum of furniture, agricultural tools, radio/TV, bike/motorbike, appliances and cell phone. Furniture includes cupboards, sofas, chairs, tables and beds. Appliances include clocks, stoves and solar panels. Value of non-land assets includes the value of livestock, durable goods, savings and the value of iron sheets used in roofing the home. Variables in PPP adjusted USD are topcoded for the highest 1 percent 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) that of large compared to small transfers. The unit of observation is the household. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 6: Treatment effects: Agricultural and Business Activities

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Wage labor primary income (dummy)	0.16	−0.00	−0.05	0.02	0.02	0.00	1372
	(0.37)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	
Own farm primary income (dummy)	0.56	−0.01	0.05	−0.00	0.00	0.02	1372
	(0.50)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	
Non-ag business primary income (dummy)	0.12	0.02	0.01	−0.02	0.01	0.01	1372
	(0.32)	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	
Non-agricultural business owner (dummy)	0.32	0.02	0.01	−0.03	0.07	0.02	1372
	(0.47)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	
Non-ag business revenue, monthly (USD)	28.62	11.15*	−2.19	5.14	12.86	−0.24	1372
	(86.25)	(5.71)	(5.51)	(10.28)	(10.76)	(8.61)	
Non-ag business flow expenses, monthly (USD)	16.61	10.18**	−0.73	6.52	9.92	−3.87	1372
	(60.12)	(4.16)	(3.77)	(7.26)	(7.75)	(6.11)	
Non-ag business profit imputed, monthly (USD)	12.01	−0.58	−1.46	−0.25	3.34	3.43	1372
	(44.10)	(3.65)	(3.21)	(6.63)	(7.90)	(5.39)	
Non-ag business profit self-reported, monthly (USD)	8.26	1.86	1.70	0.34	1.30	0.10	1372
	(24.73)	(1.73)	(1.77)	(3.21)	(3.09)	(2.83)	
Non-ag business investment in durables, monthly (USD)	0.17	0.24***	−0.10	−0.15	0.01	−0.15	1372
	(0.74)	(0.08)	(0.07)	(0.17)	(0.17)	(0.13)	
Farm revenue, monthly (USD)	9.66	0.23	−0.21	−0.10	−0.03	0.02	1372
	(8.89)	(0.54)	(0.76)	(0.90)	(0.91)	(0.82)	
Farm flow expenses, monthly (USD)	5.01	1.47***	−0.61	−0.67	−0.18	1.18*	1372
	(5.84)	(0.36)	(0.60)	(0.63)	(0.60)	(0.66)	
Farm profit, monthly (USD)	4.65	−1.21***	0.41	0.48	0.18	−1.22	1372
	(7.47)	(0.47)	(0.67)	(0.80)	(0.79)	(0.77)	
Livestock flow revenue, monthly (USD)	6.44	3.02***	−2.09**	1.24	3.70*	−1.08	1372
	(14.04)	(0.98)	(1.05)	(1.86)	(1.91)	(1.52)	
Livestock flow expenses, monthly (USD)	2.33	1.31***	−0.51	−0.49	−0.81	2.41***	1372
	(4.64)	(0.33)	(0.36)	(0.63)	(0.52)	(0.67)	
Livestock flow profit, monthly (USD)	4.11	1.68*	−1.57	1.52	4.32**	−3.60***	1372
	(13.21)	(0.94)	(1.03)	(1.77)	(1.84)	(1.37)	
Livestock sales and meat revenue, monthly (USD)	4.25	2.21***	−0.75	−0.42	−0.18	−0.54	1372
	(8.40)	(0.61)	(0.57)	(1.21)	(1.16)	(1.07)	
Total revenue, monthly (USD)	48.98	16.64***	−5.23	5.30	16.20	−1.64	1372
	(90.52)	(5.93)	(5.67)	(10.61)	(11.11)	(8.96)	
Total expenses, monthly (USD)	23.95	12.90***	−1.85	5.38	8.98	−0.18	1372
	(61.71)	(4.23)	(3.91)	(7.38)	(7.84)	(6.29)	
Total profit, monthly (USD)	20.78	−0.16	−2.63	1.37	7.60	−1.54	1372
	(46.22)	(3.74)	(3.40)	(6.72)	(8.00)	(5.39)	
Joint test (p-value)		0.00***	0.11	0.85	0.52	0.04**	

Notes: OLS estimates of treatment and spillover effects. Outcome variables are listed on the left. Variables indicating primary source of income and operation of non-agricultural enterprise are dummy variables. All other outcome variables are reported as PPP adjusted USD, converted to a monthly time horizon. Enterprise flow expenses includes the total cost of electricity, wages, water, transport, inputs, inventory and any other expenses for all enterprises owned and operated (partially or fully) by the respondent. Farm revenue includes the value of crops sold or consumed in the short and long rain seasons, expressed on a monthly basis. Farm flow expenses include seeds, fertilizers, herbicides, rental of machines, water, labor and other expenses related to agricultural production. Livestock flow revenue includes the value of animal products sold or consumed, including milk, eggs and other animal products. Livestock flow expenses include all expenses in caring for animals, such as fodder or veterinary care. Total revenue, expenses and profits are the sum across farm, animal and non-agricultural enterprises. Variables in PPP adjusted USD are topcoded for the highest 1 percent 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) that of large compared to small transfers. The unit of observation is the household. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 7: Treatment effects: Psychological well-being

	(1) Control mean (SD)	(2) Treatment effect	(3) Spillover effect	(4) Female recipient	(5) Monthly transfer	(6) Large transfer	(7) N
Log cortisol (no controls)	2.46	0.04	−0.09	−0.22**	0.26**	−0.12	2102
	(0.89)	(0.06)	(0.06)	(0.09)	(0.10)	(0.09)	
Log cortisol (with controls)	−0.03	0.05	−0.08	−0.21**	0.27***	−0.16*	2102
	(0.88)	(0.06)	(0.06)	(0.09)	(0.10)	(0.09)	
Depression (CESD)	26.48	−0.99*	−0.73	−0.98	−1.36	−1.76**	2140
	(9.31)	(0.55)	(0.78)	(0.82)	(0.89)	(0.82)	
Worries	0.00	−0.09	0.03	−0.15*	−0.13	−0.11	2140
	(1.00)	(0.06)	(0.08)	(0.08)	(0.10)	(0.09)	
Stress (Cohen)	0.00	−0.14**	0.05	0.04	0.02	−0.30***	2140
	(1.00)	(0.06)	(0.07)	(0.09)	(0.10)	(0.09)	
Happiness (WVS)	−0.00	0.18***	0.13	−0.01	0.01	0.07	2140
	(1.00)	(0.06)	(0.08)	(0.09)	(0.10)	(0.09)	
Life satisfaction (WVS)	−0.00	0.15***	0.04	−0.01	0.00	0.17**	2140
	(1.00)	(0.05)	(0.08)	(0.08)	(0.09)	(0.08)	
Trust (WVS)	−0.00	0.06	−0.09	0.08	0.06	−0.15	2140
	(1.00)	(0.06)	(0.07)	(0.09)	(0.10)	(0.10)	
Locus of control	0.00	0.03	−0.06	0.04	0.03	0.02	2140
	(1.00)	(0.06)	(0.08)	(0.10)	(0.10)	(0.10)	
Optimism (Scheier)	−0.00	0.10	0.14*	0.12	0.05	0.13	2140
	(1.00)	(0.06)	(0.07)	(0.09)	(0.10)	(0.10)	
Self-esteem (Rosenberg)	0.00	0.01	−0.11	0.23**	0.09	−0.10	2140
	(1.00)	(0.06)	(0.08)	(0.10)	(0.11)	(0.12)	
Psychological well-being index	−0.00	0.20***	0.08	0.16*	−0.10	0.35***	2140
	(1.00)	(0.06)	(0.07)	(0.09)	(0.10)	(0.10)	
Joint test (*p*-value)		0.01***	0.31	0.11	0.26	0.00***	

Notes: OLS estimates of treatment and spillover effects. Outcome variables are listed on the left. All variables are coded in z-score units, except raw cortisol, which is coded in nmol/l. 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, i.e. comparing treatment households to control households within villages. Column (3) reports the spillover effect, i.e. the treatment effect on spillover households compared to pure control households. Column (4) reports the relative treatment effect of transferring to the female compared to the male; column (5) the relative effect of monthly compared to lump-sum transfers; and column (6) that of large compared to small transfers. The unit of observation is the individual. All columns except the spillover regressions include village-level fixed effects, control for baseline outcomes, and cluster standard errors at the household level (in the spillover column, at the village level). The last row shows joint significance of the coefficients in the corresponding column from SUR estimation. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

7 Appendix

7.1 Accounting for multiple inference

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

We have catalogued below the primary groups of outcomes that we intend to consider in the analysis outlined above. For each of these outcome groups, we will construct indices (where possible) and for each of the components of these indices, and will report both unadjusted p-values as well as p-values corrected for multiple comparisons using the Family-Wise Error Rate.

7.2 Construction of indices

To keep the number of outcome variables low and thus allow for greater statistical power even after adjusting p-values for multiple inference, we construct indices for several of our groups of outcome variables. To this end, we follow the procedure proposed by Anderson (2008), which is reproduced below:

First, for each outcome variable y_{jk}, where j indexes the outcome group and k indexes variables within outcome groups, we re-code the variable such that high values correspond to positive outcomes.

We then compute the covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{j}}$ for outcomes in outcome group j, which consists of elements:

$$\hat{\Sigma}_{jmn} = \sum_{i=1}^{N_{jmn}} \frac{y_{ijm} - \bar{y}_{jm}}{\sigma^y_{jm}} \frac{y_{ijn} - \bar{y}_{jn}}{\sigma^y_{jn}} \tag{13}$$

Here, N_{jmn} is the number of non-missing observations for outcomes m and n in outcome group j, $\bar{y}_{jm}$ and $\bar{y}_{jn}$ are the means for outcomes m and n, respectively, in outcome group j, and σ^y_{jm} and σ^y_{jn} are the standard deviations in the pure control group for the same outcomes.

Next, we invert the covariance matrix, and define weight w_{jk} for each outcome k in outcome group j by summing the entries in the row of the inverted covariance matrix corresponding to that outcome:

$$\hat{\mathbf{\Sigma}}_{\mathbf{j}}^{\mathbf{-1}} = \begin{bmatrix} c_{j11} & c_{j12} & \cdots & c_{j1K} \\ c_{j21} & c_{j22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots \\ c_{jK1} & \vdots & \ddots & c_{jKK} \end{bmatrix} \tag{14}$$

$$w_{jk} = \sum_{l=1}^{K_j} c_{jkl} \tag{15}$$

Here, K_j is the total number of outcome variables in outcome group j. Finally, we transform each outcome variable by subtracting its mean and dividing by the control group standard deviation, and then weighting it with the weights obtained as described above. We denote the result $\hat{y}_{ij}$ because this transformation yields a generalized least squares estimator (Anderson 2008).

$$\hat{y}_{ij} = \left(\sum_{k \in \mathbb{K}_{ij}} w_{jk} \right)^{-1} \sum_{k \in \mathbb{K}_{ij}} w_{jk} \frac{y_{ijk} - \bar{y}_{jk}}{\sigma^y_{jk}} \tag{16}$$

Here, $\mathbb{K}_{ij}$ denotes the set of non-missing outcomes for observation i in outcome group j.

7.3 Family-wise Error Rate

Because combining individual outcome variables in indices as described above still leaves us with multiple outcome variables (viz. separate index variables for health, education, etc.), we additionally adjust the p-values of our coefficients of interest for multiple statistical inference. These coefficients are those on the treatment dummies in the basic specifications, or those on the dummies for individual treatment arms. To this end, we proceed as follows, reproduced again from Anderson (2008). A similar procedure is described in Lee & Shaikh (2013) and Romano & Wolf (2005).

First, we compute naïve p-values for all index variables $\hat{y}_j$ of our j main outcome groups, and sort these p-values in ascending order, i.e. such that $p_1 < p_2 < \cdots < p_J$.

Second, we follow Anderson's (2008) variant of Efron & Tibshirani's (1993) non-parametric permutation test: for each index variable $\hat{y}_j$ of our j main outcome groups, we randomly permute the treatment assignments across the entire sample, and estimate the model of interest to obtain the p-value for the coefficient of interest. We enforce monotonicity in the resulting vector of p-values $[p^*_1, p^*_2, \cdots p^*_J]'$ by computing $p^{**}_r = \min\{p^*_r, p^*_{r+1}, \cdots p^*_J\}$, where r is the position of the outcome in the vector of naïve p-values.

We then repeat this procedure 10,000 times. The non-parametric p-value, p_r^{fwer*}, for each outcome is the fraction of iterations on which the simulated p-value is smaller than the observed p-value. Finally we enforce monotonicity again: $p_r^{fwer} = \min\{p_r^{fwer*}, p_{r+1}^{fwer*}, \cdots p_J^{fwer*}\}$. This yields the final vector of family-wise error-rate corrected p-values. We will report both these p-values and the naïve p-values. Within outcome groups, we report naïve p-values for individual outcome variables other than the indices.