General Equilibrium Effects of Cash Transfers in Kenya: Note for Givewell

Johannes Haushofer, Edward Miguel, Paul Niehaus and Michael Walker*

October 2, 2018

Abstract

This research note provides estimates on the effects of unconditional cash transfers on subjective well-being, household asset ownership and expenditures, along with market prices, from the General Equilibrium Effects of Cash Transfers (GE) project. The GE project is a large-scale, two-level randomized controlled trial of unconditional cash transfers to poor households in rural Kenya delivered by the NGO *GiveDirectly*. Villages were randomly assigned to receive the cash transfer program, with all poor households meeting a basic means test receiving transfers in treatment villages and received roughly USD1,000 (nominal). Groups of villages were randomly assigned to high or low treatment saturation, with 1/3 of villages in low saturation areas being assigned to treatment and 2/3 of villages in high saturation areas assigned to treatment. A total of about USD 11 million was distributed across 653 treatment and control villages. Ten months after the end of transfers (on average), we find recipient households experience increases in subjective well-being (0.09 standard deviations), asset ownership and expenditure; all effects are statistically significant at the 1% level. We find a statistically significant increase in subjective well-being for ineligible non-recipient households living in treatment villages, and marginally significant effects on household expenditure.

*Haushofer: Princeton University, NBER, and Busara Center for Behavioral Economics, haushofer@princeton.edu. Miguel: UC Berkeley, emiguel@berkeley.edu. Niehaus: UC San Diego, pniehaus@ucsd.edu. Walker: UC Berkeley, mwwalker@berkeley.edu. We thank Justin Abraham, Aakash Bhalotia, Christina Brown, Genevieve Deneoux, Dennis Egger, Rachel Pizatella-Haswell, Francis Wong, and Zenan Wang for excellent research assistance, *GiveDirectly* for collaboration, the Busara Center for Behavioral Economics for survey development and piloting, and IPA-Kenya for data collection. This work has been funded by the Private Enterprise Development in Low-Income Countries (PEDL) initiative, the International Growth Centre, the Weiss Family Foundation, and an anonymous donor. Niehaus is a co-founder and president of *GiveDirectly*. This study is registered in the AEA Trial Registry as AEARCTR-0000505, `https://www.socialscienceregistry.org/trials/505`

Cross-village spillover effects are generally small and not statistically significant. Price effects also appear small, both during and after the distribution of cash.

1 Introduction

We conduct a large-scale randomized controlled trial of the unconditional cash transfer program of the NGO *GiveDirectly* (GD), which makes large unconditional cash transfers to poor households in Kenya. The magnitude of the transfers is large, around USD 1,000 (nominal) per household, or about USD 1,870 PPP, roughly 75% of annual expenditure for recipient households.[1] At the time of this study, GD targeted households living in homes with grass-thatched roofs, a basic means-test for poverty; we find 33% of households eligible in our study area. (GD currently uses a variety of targeting criteria). Treatment assignment is randomized at the village level, and within treatment villages, all households meeting GD's eligibility requirement receive the unconditional cash transfer. A second level of randomization provides variation in treatment intensity: sublocations, an administrative unit directly above the village level comprising of an average of ten villages, were randomly assigned to high or low saturation status. In high saturation sublocations, two-thirds of villages were assigned to treatment, while in low saturation sublocations, only one-third of villages were assigned to treatment.[2] The intervention involves roughly USD 11 million in transfers across 653 villages (328 treatment and 325 control villages) in one Kenyan county.

In this research note, we summarize results on subjective well-being, household assets and expenditure from our first endline survey, conducted on average 10 months after the distribution of the final transfer.[3] Subjective well-being provides a measure of household's overall welfare, while looking at asset ownership and expenditure provides some insights into how households are spending and saving their transfers. Together, these outcomes provide a sense of the welfare impacts of these transfers. It is important to note (especially if one plans to use these with GiveWell's cost-benefit calculator) that the effects on consumption expenditure are measured several months after the transfers, and therefore do not capture increased consumption in the first months after receiving the transfer; these impacts are therefore lower bounds. For household-level outcomes, we estimate the direct effects of GD's

1. Purchasing power parity values are calculated based on the World Bank's PPP conversion factor for private consumption of 46.49 for 2016, the year in which most of our endline data collection was carried out.

2. More details can be found in section 2 on the experimental design.

3. These outcomes were selected given GiveWell's specific interest in these for evaluating the effectiveness of GD's program in our study area in rural Kenya. In addition to these outcomes, our pre-analysis plan also pre-specified total household income, total business revenue, a health index, an education index, a female empowerment index, a food security index and hours worked as primary outcomes of interest.

program by comparing eligible households in treatment versus control villages, as well as spillover effects on i) ineligible households within treatment villages (within-village spillovers) and ii) saturation effects (cross-village spillovers), on average and separately for eligible and ineligible households.[4]

A key concern about cash transfers is that they may lead to inflation, diminishing the real value of the transfer and causing negative spillovers for non-recipient households. We therefore also report on price effects from 2 years of market survey data, collected as transfers were distributed through our endline surveys. Market surveys were conducted monthly from the universe of weekly markets in our study area (a total of 61 markets). The random variation in treatment intensity around market centers allows us to estimate a causal effect of cash transfers on prices within radius bands around treatment villages.[5]

Our household surveys were conducted an average of ten months after the last transfer, though there is variation in this timing (see Figure 1, Panel B and Figure 2).[6] This timing is especially important to keep in mind when considering our measure of household expenditure, as we construct an annualized measure from recall periods over the last week, last month, and last year. In particular, we measure food consumption over the last week, frequent non-food purchases over the last month, infrequent purchases over the last 12 months, and temptation goods over a combination of the last month and last week.[7] From an impact evaluation perspective, this allows us to measure whether there is a degree of persistence to household effects. However, as noted above, this approach potentially underestimates the treatment effect on these outcomes, because expenditure is likely to increase especially strongly in the first months after transfer receipt. Some households will also have received all or part of their transfers outside of the last 12 months captured in the infrequent purchases recall period as well. To provide context on the magnitudes of our results, for monetary outcomes we report coefficient estimates and we calculate effects as a percentage of the mean for households in control villages in low saturation sublocations, and as a share of the total transfer value. These timing considerations are especially important to keep in mind when interpreting results as a share of the transfer value, as households may not have received the full transfer value in the last 12 months. In addition, note that households may also generate additional income as a result of the transfers, which could also lead to increased expenditure and asset purchases.

4. Our household outcome construction and regression equations were pre-specified as part of Haushofer et al. (2017).

5. Market price outcomes and regressions are defined as part of Haushofer et al. (2016).

6. We find no significant differences in survey timing by treatment status.

7. Appendix A.3 provides more detail on the types of expenditure in each recall category.

We find that, on average ten months after the end of transfers, recipient households experience a 0.09 standard deviation increase in subjective well-being ($p < 0.01$). We find large and significant increases in reported life satisfaction by recipient households, while point estimates for depression and stress go in the direction of declines for both but are not statistically significant. In addition to gains for recipient households, we also find that ineligible households in treatment villages also report significant increases in subjective well-being of 0.11 SD ($p < 0.01$). For ineligible households, we find statistically significant effects in reported happiness and a reduction in stress.

We also find strong direct effects on household asset ownership and expenditure. On assets, recipient households report an increase of USD 177 PPP in total movable assets net of loans (this excludes land an home value), a 26 percent increase relative to the control, low saturation mean. Recipient households also report an increase in their home value of USD 359 PPP in the reported home value, a 60 percent increase relative to the control mean.[8] Taken together, the value of these asset increases corresponds to about 30 percent of the full cash transfer value (9 percent for movable assets and 19 percent for home value). Household expenditure for recipient households goes up by USD 314 PPP, a 13 percent increase relative to the control mean (17 percent of the full transfer value). All of these effects are statistically significant at the 1% level. As with other studies (see Evans and Popova 2014, for a review), we do not find an increase in temptation good expenditure. In terms of within-village spillovers, we find marginally significant positive increases in total expenditure for ineligible households in treatment versus control villages. We find no effects on asset ownership for this group. Importantly, we do not find evidence for strong negative effects on assets either, as point estimates are positive, not statistically different from zero and generally small in magnitude.

For all outcomes, cross-village spillover effects of differences in treatment saturation are generally small and not statistically significant. This holds both when looking at averages across all household types, as well as when looking at specific household types (for instance, eligible households in treatment villages).

One concern about the asset and expenditure results presented here is that they may be overstating the real magnitude of effects if the transfers lead to an increase in price inflation. This does not seem to be the case: our market price data indicates small price effects, both during and after the distribution of cash, for areas (and periods of time) receiving greater amounts of cash. Our preferred regression specification suggests there were no statistically

8. Measuring home values are difficult in contexts with limited land and home sales. We asked respondents to estimate the cost of materials and labor to build a home like their own.

significant price increases on average, and the point estimate remains small and precise. Implied contemporaneous average treatment effects on prices, both overall and for major components of the price index, are essentially zero. These results compare markets in our study area to one another; we cannot fully rule out a level shift in all markets in the study area. However, as we find limited effects across different spatial distances, this is suggestive that there was not a shift up in all markets in the area.

Taken together, these short-run results suggest that recipient households do indeed benefit from GD's program, and do not provide evidence for negative spillovers to non-recipient households, at least on the dimensions examined here. (Table 5 provides a summary of our household-level results for direct, within-village, and cross-village effects.)

These results provide evidence on GD's typical lump sum transfer procedure at the time of our study start, in 2014. GD's program for distributing lump sum transfers has continued to evolve since its implementation in our study area. At the time of this study, GD targeted poor areas (as they continue to do) and sought to make transfers to the poorest 30 to 40 percent of households in treatment villages. In the case of this study, the targeting rule is also observable by households, and (anecdotally) became well-known in our study area. One important difference is that GD now frequently does saturation targeting within villages, targeting all (or almost all) households within treatment villages for transfers. Second, for our study, transfers were distributed via mobile money in three tranches over over the course of 8 months. Recipient households now receive two payments, two months apart. In addition to differences that may arise from implementing GD's program in different contexts, these programmatic changes may also have implications for the type of effects we would expect to see if re-evaluating the program.

The rest of this note is structured as follows: Section 2 provides the study design, including information on the study context, cash transfer intervention, experimental design and data collection. Section 3 describes our main empirical specifications for both households and prices. Section 4 provides results on subjective well-being, assets, expenditure, and prices, respectively. Section 5 provides a brief discussion and concluding thoughts.

2 Study Design

2.1 Setting

This study takes place in Siaya County, Kenya, a rural area in western Kenya bordering Lake Victoria. Siaya County is predominatly Luo, the second largest ethnic group in Kenya. GD selected both Siaya County and our study region within Siaya County based on its high poverty levels.[9] GD identified target villages for expansion; in practice, these were all villages within the region that a) were not located in peri-urban areas and b) were not part of a previous GD campaign. This gives a final sample of 653 villages, spread across 84 administrative sublocations (the unit above a village), and 3 constituencies.[10]

2.2 Cash Transfer Intervention

GD provides unconditional cash transfers to poor households in rural Kenya, targeting (for villages in our study) households living in homes with thatched roofs, a basic means-test for poverty. In treatment villages, GD enrolls all households in treatment villages meeting its thatched-roof eligibility criteria ("eligible" households); approximately one-third of all households are eligible. No households in control villages receive transfers. Eligible households enrolled in GD's program receive a series of 3 transfers totaling about USD 1,000 (nominal), or USD 1,870 PPP.[11] Transfers are distributed via the mobile money system M-Pesa.[12] This is a one-time program and no additional financial assistance is provided to these households after their final large transfer.

GD's enrollment process in treatment villages consists of the following 6 steps:

1. Village meeting (*baraza*): Before beginning work in a village, GD holds a meeting of all

9. This selection was based on the 2009 Population Census, which occurred prior to devolution and the creation of county governments. Based on 2009 administrative boundaries, the study area consists of 5 of the 7 divisions in Siaya District: Boro, Karemo, Ugunja, Ukwala and Uranga. The 2009 census lists enumeration areas, which we refer to as villages.

10. 5 villages were dropped after randomization: 4 villages, all of which contained the "Town" in the name, were dropped for being too urban for GD to work in. 1 of these was assigned to treatment, the remaining were assigned to control. The boundaries of one control village were unable to be determined by field staff despite repeated efforts. This was an enumeration area created for the 2009 census that did not correspond to existing village boundaries.

11. The total transfer amount is 87,000 Kenyan Shillings (KES). The average exchange rate from 9/1/14 to 4/30/16 was 97 KES/USD. The World Bank's purchasing power parity conversion factor for private consumption for 2016, the year in which endline data was collected, is 46.49. We use this amount when converting to PPP values.

12. For more information on M-Pesa, see Mbiti and Weil (2015) and Jack and Suri (2011).

households in the village to inform villagers that GD will be working in their village, explain their program and GD as an organization. To prevent gaming, the eligibility criteria are not disclosed.

2. Census: GD staff conduct a household census of the village, collecting information on household names, contact information and housing materials. The information on housing materials are used to determine program eligibility.

3. Registration: Households identified as eligible based on the household census are visited by the registration team. GD staff confirm the eligibility of the household, inform the household of their eligibility for the program and register the household for the program. This is the point at which households learn they will be receiving transfers, as well as the amount of the transfers, the transfer schedule, and the fact that the transfer is unconditional.[13] Households are instructed to register for M-Pesa, a prerequisite for receiving the transfer. Households that do not have a mobile phone are given the option to purchase one from GD staff, the cost of which is deducted from the transfer amount.

4. Backcheck: All registered households are backchecked to confirm eligibility in advance of the transfers going out. This is an additional step to prevent gaming by households and field staff, as the census, registration and backcheck teams consist of separate staff members.

5. Transfers: The cash is transferred in a series of three payments via M-Pesa according to the following schedule: (i) the token transfer of about USD 70 nominal / USD 150 PPP ensures the system is working properly; (ii) two months afterwards, the first lump sum transfer of about USD 415 nominal / USD 860 PPP is distributed; (iii) six months after this, the second and final lump sum transfer of USD 415 nominal / USD 860 PPP is sent. If households elected to receive a mobile phone from GD, the cost of this is taken out of the second lump sum transfer. Transfers are typically sent at one time per month to all households scheduled to receive transfers in a given month.

6. Follow-up: After transfers go out, GD staff follow up via phone with transfer recipients to ensure no problems have arisen. In addition, there is a GD help line that recipients can contact. If GD staff learn that household conflicts have arisen as a result of the transfers, transfers were sometimes delayed while these problems were worked out.

13. To emphasize the unconditional nature of the transfer, households are provided a brochure with many potential uses of the transfer.

2.3 Randomization

We use a two-level randomization in order to generate variation that can be used to identify direct, within-village and cross-village spillover effects. Figure 1, Panel A presents this design graphically. We randomly assigned sublocations (or in some cases, groups of sublocations) to high or low saturation status. Then, within high saturation groups, we assigned 2/3 of villages to treatment status, while within low saturation groups, we assigned 1/3 of villages to treatment status. As noted above, within treatment villages, all eligible households receive a cash transfer, while ineligible households and households in control villages do not receive transfers.

The randomization was conducted in two batches based on GD's expansion plans. The first batch included villages in Alego constituency, where GD had previously worked. In Alego, we sought to create saturation groups in which the number of villages in our study was a multiple of 3, if it was possible to combine contiguous sublocations; this also ensured at least 3 villages were in a saturation group. We created 23 saturation groups out of a total of 39 sublocations in Siaya, 11 of which matched directly to a single sublocation. Saturation groups in Siaya had on average 10 villages. We stratified assignment of high and low saturation by the level of exposure within the saturation group (the share of villages involved in a previous GD campaign), splitting the exposure level at the median. We then randomly assigned villages to 3 groups, and randomly assigned these groups to either a) always treatment, b) treatment in high saturation, control in low saturation and c) always control. We randomly generated an order for GD to work in by first randomly ordering the saturation groups and then villages within saturation groups.

The second batch included villages in Ugunja and Ugenya constituencies. GD had not previously worked in any villages in these constituencies, so we did not stratify on any variables for these villages. Given the larger number of villages per sublocation, we also took the sublocation to be the saturation group. We assigned villages to one of three groups, pooled the "residual" villages that were not a multiple of 3, and randomly assigned 1/3 of these to the always treatment group, 1/3 to the treatment in high saturation sublocation group, and 1/3 to the always control group. GD worked first in Ugunja and then Ugenya. We generated a random order within these constituencies by first ordering locations (the administrative unit above the sublocation), then sublocations within the location, then villages within the location. Ordering based on location was used in an attempt to limit gaming by households.

Due to the large number of villages and households involved in the study, GD worked on a rolling basis across villages in the study area following the random order described above. The timing of transfers to eligible households within a village may vary for several reasons. GD generally began sending transfers to eligible households within a village once 50% of the eligible households (as identified via the census) completed the enrollment process. Villages that were above this threshold but in which GD was still working on completing the enrollment of other households would see a difference in the timing of transfers to households. If households delayed in signing up for M-Pesa, this would also introduce delays in their transfers and differences across villages. If households reported issues arising due to the transfers (such as marital problems or other conflicts), transfers may be delayed while these problems are worked out.

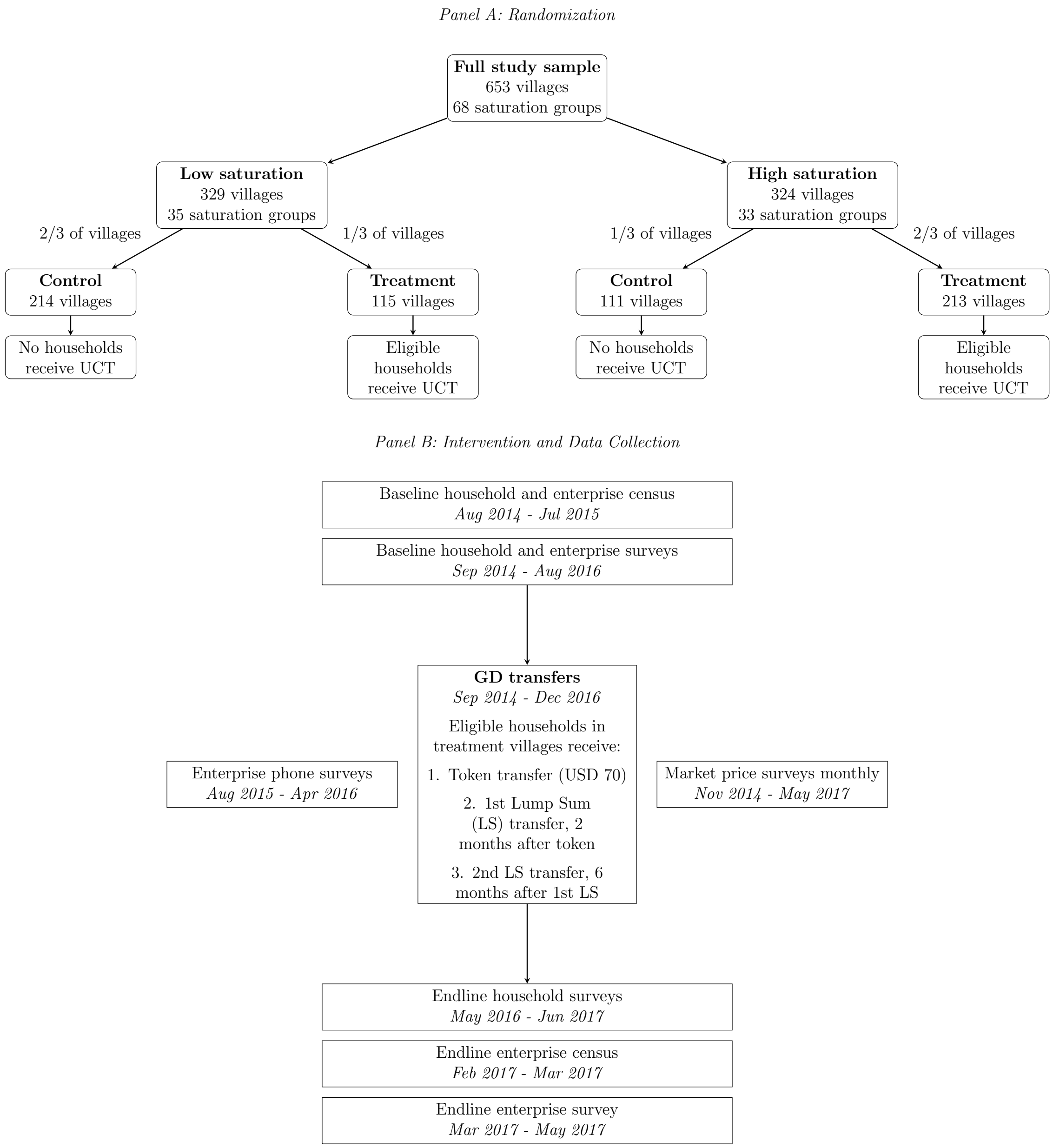
2.4 Household Data Collection

The primary data source for this analysis are household surveys. Several aspects of our data collection effort stand out: first, by including ineligible households, we can test for within-village spillover effects onto these households. Second, we tracked households that migrated from their original homestead, providing high tracking rates and allowing us to measure effects even for households that use the cash transfer to migrate. We first describe the manner in which we collected this data; Sections 2.4.1 and 2.4.2 go into greater detail on the construction of specific outcomes from these data. Figure 1, Panel B presents the calendar timeline of data collection. Figure 2 presents the timeline in reference to baseline surveys.

In advance of the distribution of transfers to a treatment village, we conducted a baseline household census and household survey. The household census was designed to be comparable to GD's census, but to ensure there was no systematic bias between their censusing methods and ours, we conducted our own censuses in all villages (both treatment and control). The census served as a sampling frame for baseline household surveys. We determined household eligibility based on the census data and targeted 12 households per village for inclusion in the study, 8 eligible households and 4 ineligible households. We randomly ordered households by treatment status, and attempted to survey the first 8 eligible and first 4 ineligible households; we refer to these households as "initially-sampled" households. For couples, we randomly selected either the male or female to be the "target" respondent; if we could not reach the target, but the spouse/partner was available, we surveyed the spouse/partner.

If an initially-sampled household was not available to be surveyed on the day we visited the village for baseline surveys, we replaced this household with the next one on the list in

Figure 1: Study design

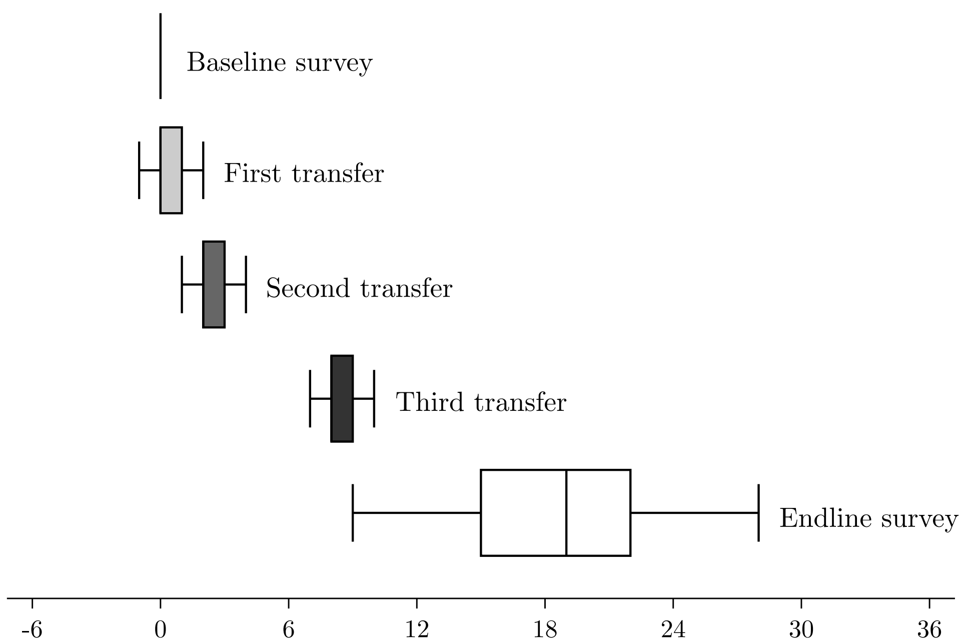


Notes: This figure outlines the two-level randomized controlled trial experimental design. 653 villages were grouped into saturation groups based on the sublocation (the administrative unit directly above the village level) in which they are located. Saturation groups are then randomly assigned to either high or low saturation status. In the 33 high saturation status groups, two-thirds of villages are assigned to treatment status, while in the 35 low saturation status groups, one-third of villages are assigned to treatment status. In the 328 treatment villages, eligible households receive an unconditional cash transfer, while no households within control villages receive a transfer. 99% of transfer value was sent by December 2016.

order to ensure that we surveyed 12 households in each village; we refer to these households as “replacement” households. Lastly, we refer to households that were initially-sampled but unable to be surveyed as “missed baseline” households.

Endline surveys target all “initially sampled” and “replacement” households. For households that were baselined, we attempt to survey the same respondent that was surveyed at baseline. Endline surveys began at the end of May 2016 and concluded in June 2017. The median survey date is about 18 months after the baseline surveys and 10 months after the second (and final) lump sum cash transfer. Figure 2 shows the variation in endline timing in our sample using a box plot for the number of months since each village’s “experimental start month,” which we define based on GD’s pace and the order in which they worked in, so that both treatment and control villages have a start month.

Figure 2: Study timeline, with respect to village experimental start month



Notes: This figure summarizes survey dates and transfer timing using box plots at the village level. The x-axis represents months since the “experimental start month” in each village. The experimental start month is assigned by evenly spacing GD’s pace across all villages in a subcounty, and allocating villages to a transfer start month based on the order in which villages were assigned to be visited. It thus provides a treatment start date for both treatment and control villages. For full details, see Haushofer et al. (2016).

2.4.1 Household subjective well-being outcomes

The subjective well-being index is a weighted, standardized index of depression, happiness, life satisfaction, and stress, appropriately signed so that positive values indicate greater subjective well-being. We follow the procedure proposed by Anderson (2008) to generate weights and normalize values, and report results for both the overall index and components of the index. Note that we do this separately for eligible and ineligible households, as we believe these moments are the most relevant for each group. Our measure of depression comes from the 10-question CES-D scale, with higher scores indicating greater levels of depression. Happiness and life satisfaction measures both come from the World Values Survey (WVS), and in both cases higher values indicate positive outcomes. Lastly, we use Cohen's 4-question Perceived Stress Scale, where higher values indicate greater stress levels. Both depression and stress are negatively coded when generating the subjective well-being index.

There are several differences between the baseline and endline subjective well-being measure, but the primary purpose is to increase precision of the estimated treatment effects, which will occur as long as the baseline measure is correlated with the endline measure. First, we did not ask the perceived stress scale as part of baseline surveys; baseline values of the subjective well-being measure omit this scale. Second, we used a 20-question CES-D scale at baseline, and include all 20 questions when calculating baseline CES-D values.

2.4.2 Household asset and expenditure outcomes

To better understand both household well-being as well as how the transfers affect household asset stocks and spending, we look at effects on households assets and expenditure. We report on two measures of household asset ownership: the value of non-land, non-home assets, plus any lending to others net of the household's borrowing, along with the household's home value. To measure home value, we ask households for the estimated cost of the labor and materials it would take to build a house similar to their own. Both outcomes were collected at baseline and endline, so we use ANCOVA specifications when analyzing these outcomes.

Our main measure of household expenditure seeks to capture total expenditure in the last 12 months, which we calculate as the sum of total food consumption in last 7 days, frequent purchases in last month and infrequent purchases in last 12 months, all annualized to yearly values. Due to survey time constraints and for cross-study consistency with Haushofer and Shapiro (2016), we did not conduct a full LSMS-style consumption / expenditure module. Instead, following Haushofer and Shapiro (2016), we collected totals on consumption categories. A full listing of all the expenditure data we collected is available in Appendix A.

As is common with many income and expenditure measures, we take the standard approach of winsorizing the top 1% of monetary variables.

2.5 Market Price Data Collection

We conducted monthly market surveys in 61 markets across Alego, Ugunja and Ukwala; these constitute all markets in our study area that have a weekly market day, the inclusion criterion. The market surveys were conducted monthly, starting around 5 days after GD sends out transfers for the month and running for the course of one week. The surveys collect price information on food products (grains, vegetables, fruit, meat), livestock, hardware, "duka" (kiosk store) products (non-food and packaged food), and other products (fuel, health items, household items, farming items) from 3 vendors of each product in each market. Vendors may or may not be the same across months.

The market surveys began in different months in each of the three subcounties based on GD's rollout. We have at least one month of pre-treatment data for all markets in each subcounty.[14] Market price surveys continued for all markets in all subcounties through January 2017.

Due to time and budget constraints with baseline survey data collection, we did not conduct a consumption/expenditure module at baseline. In order to have a measure of household expenditure unaffected by the cash transfers, we instead make use of expenditure shares collected in the vicinity of our study collected as part of the Kenya Life Panel Survey. The Kenya Life Panel Survey (KLPS) is a longitudinal dataset of nearly 10,000 Kenyan youth that participated in one of two previous randomized NGO programs – one which provided merit scholarships to upper primary school girls in 2001 and 2002 (Kremer, Miguel, and Thornton 2009), and one which provided deworming medication to primary school students during 1998-2002 (Miguel and Kremer 2004) in an area of western Kenya near our study. We use KLPS-3 survey data as a baseline for expenditure shares. We prefer using existing expenditure data for our main specifications rather than endline data because changes in expenditure shares could be endogenous.[15]

14. In Alego, this first month is retrospective; the first round of market surveys was conducted in November 2014, immediately after the first households treated by GD in Alego received their first large transfer, but we asked about prices each month back to August 2014, the last month before any households in our study had received their first, token transfers from GD. In Ugunja and Ukwala, price surveys began in January 2015 and March 2015, respectively, the last month before any households in these subcounties received their token transfer from GD, and also collected three months of retrospective data.

15. This follows the approach we pre-specified as part of Haushofer et al. (2016).

3 Empirical Strategy

We first turn our attention to estimating direct treatment effects for eligible households (transfer recipients). We then outline our strategy for measuring externality effects, both within-village externality effects for ineligible households in treatment versus control villages, and cross-village externalities, making use of variation in treatment saturation intensity at the sublocation level.

3.1 Direct treatment effects for household outcomes

We use data from all eligible households – "initially-sampled" households (both those that were baselined and missed at baseline) and "replacement" households – as part of our main specifications. We base our classification of eligible households on GE household census data. This is analogous to an intention-to-treat (ITT) analysis.

We collected information on subjective well-being and asset ownership at baseline, so our main specification for these outcomes is an ANCOVA that conditions on baseline values of the outcome variable; for households that we missed at baseline, we include an indicator that the household was missed at baseline, and include the mean value of the baseline variable in the regression equation.

$$y_{ihvs,t=1} = \beta_0 + \beta_1 T_{vs} + \beta_2 H_s + \delta_1 y_{ihvs,t=0} + \delta_2 M_{ihvs} + \varepsilon_{ihvs} \tag{1}$$

Here, h indexes the household, v indexes the village, s indexes the sublocation and t indicates whether the variable was measured at baseline or endline. As subjective well-being is measured for the survey respondent, i indexes the household member that we surveyed to collect this information (this may or may not be the individual listed as the recipient by GD). T_{vs} is an indicator for households residing in a treated village. β_1 identifies the effect of receiving a transfer compared to eligible households in control villages. H_s is an indicator for living in a high-saturation sublocation, which we control for as it was part of the research design, but we do not report β_2 as part of our main tables on direct effects; we turn to estimating saturation effects in the next section. Following McKenzie (2012), we condition on the baseline values of the outcome variable $y_{hv,t=0}$ to improve statistical power. When $y_{hv,t=0}$ is missing for an observation, we include an indicator term for missingness M_{ihvs} and replace $y_{hv,t=0}$ with its mean. Where we do not have baseline data (such as for household expenditure), we estimate equation (1) excluding the baseline terms. We cluster standard errors at the village level, our unit of randomization for treatment status. This provides the most precise estimate of the direct treatment effect, the coefficient on β_1, the only coefficient we report as part of our main tables.

3.2 Externality effects for household outcomes

To estimate within-village externalities on ineligible households, we restrict our sample to ineligible households and use Equation (1). All variables and standard errors are the same as for direct effects, though one should expect standard errors to increase as a) we have fewer ineligible relative to eligible households (roughly half the sample size) and b) we would expect greater variation due to greater differences in wealth and income across ineligible compared to eligible households.

To measure cross-village externalities, we pool data from all households and include interaction terms between households' eligibility status, village treatment status, and sublocation saturation status:

$$\begin{aligned} y_{ihvs,t=1} = & \beta_0 + \beta_1 T_{vs} + \beta_2 E_{hvs} + \beta_3 H_s + \beta_4 T_{vs} \times E_{hvs} + \beta_5 T_{vs} \times H_s \\ & + \beta_6 E_{hvs} \times H_s + \beta_7 T_{vs} \times E_{hvs} \times H_s + \delta_1 y_{ihvs,t=0} + \delta_2 M_{ihvs} + \varepsilon_{ihvs} \end{aligned} \tag{2}$$

$y_{ihvs,t=1}$ again is our outcome of interest, measured either at an individual or household level. T_{vs} is an indicator for households residing in a treated village, E_{hvs} is an indicator for whether the household is eligible for transfers, and H_s is an indicator for living in a high-saturation sublocation. When baseline data is available, we include it in the regression (terms $y_{ihvs,t=0}$ and M_{ihvs}). We cluster standard errors at the saturation group level to estimate the effects of saturation.

This specification allows us to estimate the following effects, clustering standard errors at the saturation group level, which we report in our saturation effects tables:

1. Treated eligible saturation regression: Effect of saturation on eligible households in treatment villages:

   $$E[y_{ihvs,t=1} \mid T = 1, E = 1, H = 1] - E[y_{ihvs,t=1} \mid T = 1, E = 1, H = 0] = \beta_3 + \beta_5 + \beta_6 + \beta_7$$

2. Untreated eligible saturation regression: Effect of saturation on eligible households in

control villages:

$$E[y_{ihvs,t=1} \mid T=0, E=1, H=1] - E[y_{ihvs,t=1} \mid T=0, E=1, H=0] = \beta_3 + \beta_6$$

3. Treated ineligible saturation regression: Effect of saturation on ineligible households in treatment villages:

$$E[y_{ihvs,t=1} \mid T=1, E=0, H=1] - E[y_{ihvs,t=1} \mid T=1, E=0, H=0] = \beta_3 + \beta_5$$

4. Untreated ineligible saturation regression: Effect of saturation on ineligible households in control villages:

$$E[y_{ihvs,t=1} \mid T=0, E=0, H=1] - E[y_{ihvs,t=1} \mid T=0, E=0, H=0] = \beta_3$$

5. Pooled saturation regression: Effect of saturation on eligible and ineligible households in treatment and control villages: Average of 1.–4., weighted by share of treatment vs. control villages and eligible vs. ineligible households.[16]

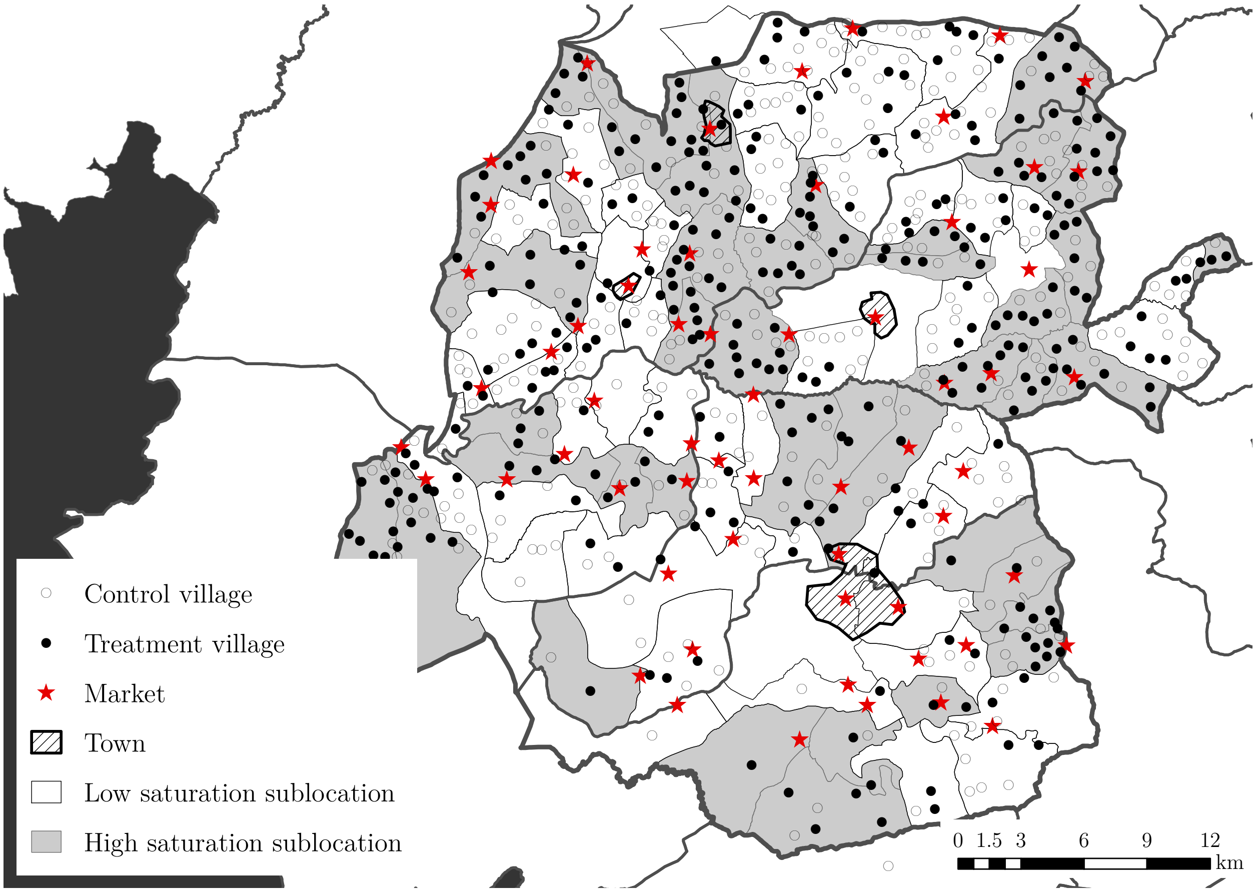
3.3 Price effects

Weekly markets play an important role in the local economy, so capturing whether prices change at the market level serves as an important factor in household welfare. Our randomized treatment assignment creates variation in the number of recipient households around each market, and hence the amount of cash being transferred into areas surrounding each market over time. This can be seen in Figure 3, which plots study villages and weekly market centers across our study area. Markets are denoted by red stars, treatment villages are denoted by solid circles, and control villages by hollow circles. Areas shaded in gray are high saturation sublocations, where one can visually see the larger share of treatment villages.

To estimate price effects, we make use of variation in the amount of cash distributed around weekly market centers in our study area. The richness of our panel data on prices allows us to control for both market and month fixed effects across all of our specifications. Controlling for market fixed effects accounts for any systematic price differences across mar-

16. Our pre-analysis plan noted that this was the main outcome of interest for measuring overall effects of saturation. Our pre-analysis plan also outlined additional linear combinations of the effects of 1.-4., pooling all (in)eligible households across treatment and control villages, or pooling all treatment households across eligibility categories. As these can all be calculated from 1.-4., we focus on these terms, and the overall pooled saturation regression calculation, as part of this research note.

Figure 3: Map of villages and markets in study area



Notes: This figure plots study villages, sublocation boundaries and weekly markets in the study area in Siaya County, Kenya. Control villages are denoted by hollow circles, treatment villages are denoted by solid circles, and red stars indicate the locations of markets. High saturation sublocations are shaded in gray, while low saturation sublocations are those in white. Town boundaries are shaded with diagonal lines.

kets, as well as differences in the share of eligible households located around markets. Month fixed effects account for seasonal differences and other time trends that affect all markets in our study area. Identification thus comes from differences over time in the amount transferred to households around a market.

A challenge is to determine the "relevant" spatial distance and temporal lags that may influence current prices. Given the lack of academic knowledge about the spatial spillover effects that may occur due to cash transfer programs, we have no a priori reason to choose a particular set of radii or temporal lags as empirically relevant. However, we do expect weak monotonicity to hold (i.e. if we do not find effects based on treatment intensity 2-4 kilometers from a market, we would not expect to find effects 4-6 kilometers from the market).

We pre-specified a procedure to select the "optimal" spatial and lagged specification as part of Haushofer et al. (2016). We use a two-step procedure:

1. Select the optimal number of spatial buffers from 0 to 20km (in 2km increments) while fixing the time dimension. For each buffer, we sum transfers over the last 3 months only, and run specifications including sequentially more buffers. This creates a series of nested models; we select the specification that minimizes the Schwartz Bayesian Information Criterion (BIC), while maintaining weak monotonicity, as the "optimal" number of buffers.

2. Including only the buffers up to the optimal distance as determined by step 1, we now use monthly data on prices and transfer amounts, and include sequentially more lags of the monthly transfers to the optimal buffers, from 0 up to 18 months. The optimal number of lags is again determined by the specification that minimizes the BIC.

Specifically, we estimate the following equation to determine relevant spatial buffers:

$$y_{mt} = \sum_{R=0-2km}^{18-20km} \beta_R \left(Amt_{mtR} + Amt_{m(t-1)R} + Amt_{m(t-2)R} \right) + \alpha_m + \lambda_t + \varepsilon_{mt} \tag{3}$$

where y_{mt} is the aggregate (log) price index of interest for market m in month t, Amt_{mtR} is the per-household amount of cash transferred within radius band r of market m in month t, which we aggregate into a measure of the amount transferred in the last quarter (i.e. months t, $t-1$, and $t-2$), α_m is a market fixed effect, λ_t is a monthly fixed effect, and ε_{mt} is an error term.[17]

Let $\bar{R}$ denote the set of selected radii. To determine if there are time effects over broader horizons, we then use the Schwartz BIC to select the optimal number of lagged monthly values to include in the regression equation:

$$y_{mt} = \sum_{r \in \bar{R}} \sum_{l=0}^{M} \beta_{rl} Amt_{m(t-l)r} + \alpha_m + \lambda_t + \varepsilon_{mt} \tag{4}$$

where we include up to the maximum number of months M that we have in our data.

When looking at components of the market price index, we repeat this algorithm for each outcome.

17. These radii bands match what we pre-specified in Haushofer et al. (2016). However, as there is very little variation in the proportion of treated eligible households after 10km, and our standard errors are similar in magnitude for 1km bands as to 2km bands, we may report results from a specification using 1km radii bands going out to 10km in the future. The results remain very similar across both specifications.

4 Results

We now outline results on subjective well-being, an overall measure of household welfare, along with household expenditure and asset ownership, in order to provide some information on how households are spending the transfer. Note again that, for most households, the recall period for expenditure does not include the all of the months in which a household received transfers, and that we did not ask households directly how they spent the transfer income. Our expenditure and asset measures provide information on households at a snapshot in time, a number of months after they have finished receiving transfers. Lastly, as one concern about monetary outcomes is that their value may be eroded by inflation, we present results on price effects.

4.1 Subjective Well-Being Results

We find statistically significant increases in subjective well-being for recipient households. Table 1, Panel A presents results on the direct effects of the GD program on eligible households in treatment versus control villages. Each row represents a separate outcome and regression, and the coefficients in column 2 for treatment village represent coefficients for β_1 in Equation (1). Recipient households report a 0.09 SD increase in their overall subjective well-being. This is driven by a strong, statistically significant increase in life satisfaction, though point estimates go in the direction of reduced depression and stress, and a slight increase in happiness. Interestingly, ineligible households in treatment villages also report large positive increases in subjective well-being (Panel B), of 0.11 SD. Here the statistically significant components are an increase in happiness and decrease in stress. As with recipient households, point estimates also indicate a decrease in depression and improved life satisfaction. As with expenditure, we do not find any statistically significant saturation effects, and point estimates are typically small, especially for the pooled saturation regression.

Taken together, this suggests that recipient households benefit in terms of increased subjective well-being, and that there are no negative spillovers (at least on average) to other households in the study area.

4.2 Assets and Expenditures

We now turn to presenting results on monetary outcomes, specifically household asset ownership and household expenditure, in order to gain insights into how (past) transfer spending

Table 1: Subjective Well-Being Results, Treatment vs Control Villages

	(1) Control, Low Sat Mean (SD)	(2) Treat	(3) Baseline controls	(4) N
Panel A: Direct effects (eligible households)				
Subjective well-being index	0.00	0.09***	Y	5,419
	(1.00)	(0.03)		
Depression (CESD 10 question)	10.66	–0.19	Y	5,419
	(5.46)	(0.16)		
Happiness	0.00	0.01	Y	5,419
	(1.00)	(0.03)		
Life satisfaction	0.00	0.15***	Y	5,419
	(1.00)	(0.03)		
Perceived stress	0.00	–0.03	N	5,418
	(1.00)	(0.03)		
Panel B: Within-village spillovers (ineligible households)				
Subjective well-being index	0.00	0.11***	Y	2,807
	(1.00)	(0.04)		
Depression (CESD 10 question)	11.38	–0.34	Y	2,807
	(5.36)	(0.23)		
Happiness	0.00	0.10**	Y	2,807
	(1.00)	(0.04)		
Life satisfaction	0.00	0.04	Y	2,805
	(1.00)	(0.04)		
Perceived stress	0.00	–0.10**	N	2,803
	(1.00)	(0.04)		

Notes: The *Subjective Well-Being Index* is a weighted, standardized average of depression, happiness, life satisfaction, and perceived stress, appropriately signed so that positive values represent better subjective well-being. Subsequent rows report endline components of the subjective well-being index; greater values of the CES-D score and Perceived Stress score correspond to lower well-being. Baseline values of the subjective well-being index do not include the Perceived stress measure, as this was not collected at baseline.

Table 2: Subjective Well-Being Estimated Effects

	(1) Pooled saturation regression	(2) Treated eligible Saturation	(3) Untreated eligible saturation	(4) Treated ineligible saturation	(5) Untreated ineligible saturation	(6) Baseline controls	(7) N
Subjective well-being index	0.02	0.03	0.05	0.01	0.02	Y	8,226
	(0.03)	(0.04)	(0.05)	(0.06)	(0.06)		
Depression (CESD)	–0.16	–0.23	–0.09	–0.01	–0.30	Y	8,226
	(0.17)	(0.20)	(0.26)	(0.35)	(0.32)		
Happiness	0.01	0.02	0.06	–0.04	0.02	Y	8,226
	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)		
Life satisfaction	0.01	–0.04	0.02	0.03	0.01	Y	8,224
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)		
Perceived stress	–0.02	–0.06	–0.05	–0.01	0.00	N	8,221
	(0.03)	(0.05)	(0.04)	(0.08)	(0.06)		

Notes: This table presents linear combinations of coefficients from Equation (2). Standard errors clustered at the saturation group level. * denotes significance at 10%, ** denotes significance at 5%, and *** denotes significance at 1%. The *Subjective Well-Being Index* is a weighted, standardized average of depression, happiness, life satisfaction, and perceived stress, appropriately signed so that positive values represent better subjective well-being.

decisions influence the value of these at the time of our endline survey. Results are shown in Tables 3 and 4. Measures of asset ownership capture the stock of household assets at the time of our survey. We focus on household movable assets (i.e. excluding home and land values) net of any borrowing and lending, as well as household home values. As previously noted, our main measure of household expenditure seeks to capture total expenditure in the last 12 months, which we calculate by annualizing total food consumption in last 7 days, frequent purchases in last month, infrequent purchases in last 12 months, and temptation good purchases in the last week (alcohol and tobacco) and last month (gambling). We thus do not measure food expenditure and frequent non-food purchases in contemporaneous transfer months. In addition, the 12 month recall period for infrequent purchases will not cover the period in which some households received their full transfer amount. We know that households do spend and invest most of the transfer amount in the months immediately following transfer receipt. Households making productive investments may also be able to increase their earnings, which could either be invested or consumed. For both assets and expenditure outcomes, we winsorize the top 1% of values by eligibility status.

We find recipient households report sizable increases in both asset ownership and expenditure ten months after transfers (Table 3, Panel A). Recipient households have USD 177 PPP higher movable assets (non-land, non-home, net of loans), a 26 percent increase, and an increase of USD 359 PPP in their home value (the cost of materials and labor to build a home like theirs - note that this still excludes land value), a 60 percent increase. The magnitude of these increases is equivalent to about 30 percent of the transfer value. Recipient households also report an increase of USD 314 PPP in total expenditure in the last 12 months, a 13 percent increase on the mean for eligible households in control villages in low saturation sublocations. The coefficient point estimate represents 17 percent of the total value of the transfer. We find significant increases for recipient households on food expenditure (6 percent), frequent non-food purchases (8 percent), and infrequent purchases (49 percent) compared to eligible households in control villages. We find no increase in temptation good spending by recipient households, consistent with other studies (Evans and Popova 2014). These results suggest that recipient households are indeed still benefiting from their transfers, as annualized expenditure for shorter recall components (food and frequent non-food purchases) remain higher, and households also have a higher stock of assets they can spend down if needed.

In Panel B of Table 3, we turn to effects on ineligible households in treatment versus control villages. While ineligible households do not report any significant increases in asset ownership (and point estimates are small), we do find that ineligible households in treatment villages report a marginally significant increase in consumption (USD 170 PPP), 9 percent of the typical transfer value and 7 percent of the control group mean. The marginally significant and positive point estimate on food consumption of USD 86 PPP accounts for over half of the increase in total expenditure for ineligible households. As with eligible households, we find no change in temptation good expenditure for ineligible households in treatment versus control villages.

Table 4 presents results on saturation effects. We do not find any statistically significant cross-village spillover effects on assets or expenditure; this generally holds both when pooling all households and estimating the effect of being in a high versus low saturation sublocation, and when looking at high versus low sublocations by treatment and eligibility categories. When pooling across all households, point estimates are generally small and statistically insignificant. Likewise, we find no statistically significant effects for (though standard errors are larger for some estimates); treated eligible households in high versus low saturation sublocations have a marginally significant decrease in temptation good expenditure, but given the numerous point estimates, we do not want to overinterpret this single number. Most importantly, these results suggest that there are not large negative spillovers either within or across villages; if anything, within-village spillovers are positive across these outcomes.

Table 3: Household Expenditure Results, Treatment vs Control Villages

	(1) Control, Low Sat Mean (SD)	(2) Treat	(3) Share of control, low sat mean (monetary outcomes)	(4) Share of transfer value (monetary outcomes)	(5) Baseline controls	(6) N
Panel A: Direct effects (eligible households)						
Household non-land, non-home assets (net loans)	679.20	177.04***	0.26	0.09	Y	5,422
	(780.37)	(22.92)				
Cost of materials and labor to build house	592.78	358.58***	0.60	0.19	Y	5,398
	(693.78)	(23.44)				
Total household expenditure, last 12 months (annualized)	2,504.81	314.17***	0.13	0.17	N	5,422
	(1,713.12)	(52.17)				
Food consumption expenditure, last 12 months (annualized from last week)	1,563.33	97.60***	0.06	0.05	N	5,420
	(1,027.47)	(31.37)				
Frequent non-food purchases, last 12 months (annualized from last month)	467.09	39.59**	0.08	0.02	N	5,419
	(534.87)	(15.47)				
Infrequent purchases, last 12 months	408.39	201.16***	0.49	0.11	N	5,422
	(823.21)	(32.68)				
Total expenditure on temptation goods, last 12 months (annualized)	4.74	0.50	0.11	0.00	N	5,419
	(13.75)	(0.41)				
Panel B: Within-village spillovers (ineligible households)						
Household non-land, non-home assets (net loans)	1,366.06	40.98	0.03	0.02	Y	2,814
	(1,555.25)	(60.21)				
		[1.00]				
Cost of materials and labor to build house	2,678.41	63.26	0.02	0.03	Y	2,793
	(5,477.85)	(177.83)				
Total household expenditure, last 12 months (annualized)	2,588.66	169.91*	0.07	0.09	N	2,815
	(2,081.95)	(91.75)				
Food consumption expenditure, last 12 months (annualized from last week)	1,627.49	86.21*	0.05	0.05	N	2,808
	(1,115.23)	(47.60)				
Frequent non-food purchases, last 12 months (annualized from last month)	447.26	24.83	0.06	0.01	N	2,808
	(614.38)	(24.83)				
Infrequent purchases, last 12 months	693.65	−29.44	−0.04	−0.02	N	2,815
	(4,072.23)	(82.74)				
Temptation good expenditure, last 12 months (annualized)	30.75	1.53	0.05	0.00	N	2,808
	(105.06)	(4.31)				

Notes: * denotes significance at 10%, ** denotes significance at 5%, and *** denotes significance at 1%. Standard errors reported in parentheses. Monetary variables are reported in USD PPP and topcoded at the 99th percentile.

Table 4: Household Expenditure Saturation Effects

	(1) Pooled saturation regression	(2) Treated eligible Saturation	(3) Untreated eligible saturation	(4) Treated ineligible saturation	(5) Untreated ineligible saturation	(6) Baseline controls	(7) N
Household non-land, non-home assets (net loans)	49.46	−23.17	28.95	30.68	114.82	Y	8,236
	(39.06)	(30.93)	(35.16)	(79.75)	(89.98)		
Cost of materials and labor to build house	−62.86	52.31	21.66	−103.48	−122.08	Y	8,191
	(113.58)	(41.41)	(34.78)	(264.50)	(222.81)		
Total household expenditure, last 12 months (annualized)	37.22	−95.23	64.13	44.73	82.48	N	8,237
	(75.34)	(86.60)	(70.55)	(138.26)	(152.12)		
Food consumption expenditure, last 12 months (annualized from last week)	16.14	11.52	50.83	−0.61	17.85	N	8,228
	(42.62)	(58.52)	(42.39)	(78.20)	(75.85)		
Frequent non-food purchases, last 12 months (annualized from last month)	1.90	−8.52	0.31	−15.67	25.49	N	8,227
	(17.53)	(23.38)	(22.81)	(32.86)	(39.52)		
Infrequent purchases, last 12 months	−36.41	−73.21	21.24	82.59	−165.83	N	8,237
	(62.51)	(55.66)	(32.99)	(64.50)	(160.53)		
Temptation good expenditure, last 12 months (annualized)	−1.46	−13.78*	−2.32	4.10	−0.43	N	8,227
	(4.04)	(8.01)	(7.57)	(7.26)	(5.35)		

Notes: This table presents linear combinations of coefficients from Equation (2). Standard errors clustered at the saturation group level. Monetary values reported are in USD PPP terms and topcoded at the 99th percentile. * denotes significance at 10%, ** denotes significance at 5%, and *** denotes significance at 1%.

Table 5: Summary of household results

	Direct Effect *Eligible HHs*	Within-Village Spillover *Ineligible HHs*	Pooled Saturation Effect *All HHs*
Household non-land, non-home assets (net loans)	26%***	3%	4%
Total household expenditure, last 12 months (annualized)	13%***	7%	1%
Subjective well-being	0.09 SD***	0.11 SD	0.02 SD

Notes: All percentages are with respect to the mean for households in control villages in low saturation sublocations for the relevant sample. * denotes significance at 10%, ** denotes significance at 5%, and *** denotes significance at 1%.

4.3 Price results

We report results from our pre-specified algorithm (PAP-Algorithm) for the overall price index and 4 sub-indices: food items, non-durable goods, durable goods and temptation goods.The algorithm selects only the closest buffer (0-2 km), and no lags for all price indices (in other words, only the concurrent month). The independent variable is thus the fraction of annual GDP (proxied by annualized per capita expenditure by control households in low saturation sublocations) distributed in cash within 0 to 2 kilometers of a market within a month.

Table 6 contains the optimal regressions for each price index as selected by the PAP-Algorithm. In general, price effects both overall and across various product classes appear small, and estimates are not statistically significant. If GD transferred 1 percent of GDP to households within 0-2 kilometers of a market in one month, we would expect the overall price index in that market to rise by 0.2 percent. This is very small, and suggests a reason why the algorithm only selects a minimal number of spatial and temporal lags: The BIC is based on how much additional variation in prices is explained by including transfers further away and further in the past. If price effects are really small, the explanatory power of additional lags is small and they are therefore not included.

To help interpretation, we also calculate the implied price effects for transfer levels (as a fraction of annual GDP) we see in our data. From the optimal regression as selected by our pre-specified algorithm, we generate two summary statistics for the effect of treatment on each price index: the average treatment effect (ATE) and the average maximum effect. The average treatment effect corresponds to the average effect of treatment on the log price index over the entire study period (September 2014 – March 2017). In other words, it is the average price effect across all markets and all month in which transfers went out to any market in the study area. We calculate this by evaluating the optimal specification at the average treatment amounts in the 0 to 2km buffer around each market over the entire study period. The average maximum effect corresponds to the average of the maximum effect of treatment on the log price index in each market. In other words, it is the predicted maximum price effect each market experiences over the study period, averaged across all markets. We first calculate the maximum effect in each market by finding the maximum amount that went to the 0 to 2km buffer around that market in any month. We then average the maximum predicted effects across all markets.

As can be seen in Table 6, both the ATE and average maximum effect are small and precisely estimated near zero across all product categories. Given our small standard errors, we can rule out large price effects: With 95% confidence, the ATE is below 0.0006 log points. For the maximum effect, the upper bound of the 95% confidence interval is 0.008 log points (or about 0.8 percentage points).

In addition to small overall price effects, we find small price effects across almost all product categories in our data, with some categories even experiencing price declines in response to greater transfer amounts as a fraction of GDP. In particular, the fact that food prices are in line with the overall price index, and that durables do not increase dramatically in price, suggests that non-recipient households are not significantly adversely affected due to price increases in response to a cash transfer program.

Table 6: Optimal regression specifications as determined by the PAP algorithm

	(1) All goods	(2) Food items	(3) Non-durables	(4) Durables	(5) Temptation goods
Fraction of annual GDP transferred in concurrent month (0 to 2km)	0.209	0.194	0.289	-0.131	-0.055
	(0.167)	(0.170)	(0.508)	(0.373)	(0.040)
Implied ATE	0.000	0.000	0.001	-0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Implied average maximum effect	0.003	0.003	0.004	0.000	0.000
	(0.002)	(0.002)	(0.007)	(0.000)	(0.000)
N	1586	1586	1586	1586	1586
R-sq	0.652	0.675	0.250	0.316	0.557

Notes: Each column represents a regression of the logarithm of a price index on the "optimal" number of lags and distance buffers of Give Directly transfers in terms of annual GDP per person in each buffer. Each regression includes a full set of market and month fixed effects. The implied ATE is calculated by evaluating the regression specification at the average level of treatment intensity between September 2014 and March 2017, the time during which transfers went out to any market in the study area. The average maximum effect is calculated by using the optimal specification to predict the largest effect across all months for each market, and then averaging this maximum effect across markets. Standard errors (in parantheses) are calculated as in Conley (1999, 2008), and we allow for spatial correlation in the errors up to 10km and for auto-correlation up to 12months. * denotes significance at 10%, ** denotes significance at 5%, and *** denotes significance at 1%..

We have explored several alternative specifications with different assumptions around spatial and temporal lags. While not exhaustive, these suggest an elasticity of price changes with respect to the fraction of GDP transferred between 0.1 and 0.2, similar to our estimates above. This implies that a program transferring 10% of GDP into an area in a month would lead to a 2% increase in prices. A transfer of this magnitude is quite large relative to many government programs, particularly in terms of the amounts distributed per month, and especially in developing countries. So, our estimates suggest that price effects are expected to be small even from very large cash transfer programs.

One caveat to these findings is that we are making comparisons within our study area. We cannot fully rule out that there was not a price increase throughout the full study area compared to neighboring areas. That said, we do not find much variation in prices across transfer amounts within our study area, despite sizeable variation in treatment intensity: The 90-10 percentile range for total amount in annual per capita GDP going out to the 0 to 2km buffer around each market over the entire period is [2%, 11%]. This is suggestive that we would not find price effects when comparing our study area to outside areas. Ongoing work seeks to compare price trends in our study are with trends in nearby areas. We also plan to address whether there are heterogeneous price effects based on market access or remoteness for markets in our study area.

5 Discussion

As with many other studies of cash transfer programs, we find that recipient households benefit, reporting higher measures of subjective well-being, and increased asset ownership and expenditure 10 months (on average) after receiving their last transfer. We see no increase in spending on alcohol, tobacco or gambling. The increase in subjective well-being among recipient households is driven in part by a strong, statistically significant increase in life satisfaction, though point estimates go in the direction of declining depression and stress.

In addition to measuring direct effects on recipient households, we can also measure spillover effects. We do not find evidence of negative within-village externalities to households that do not meet GD's eligibility criteria. If anything, these households are slightly better off, reporting significantly higher subjective well-being and marginally significantly higher expenditure. We also do not find strong evidence for cross-village spillover effects on expenditure and subjective well-being, both overall and for each treatment by eligibility subgroup.

Part of the reason for the small spillover effects may be driven by the fact that overall, price effects appear to be small, estimated as a 0.2% increase in contemporaneous prices for each 1% of GDP distributed within a highly-localized (2 km buffer) of the weekly market, shortly after transfers were distributed. Even during months in which a large amount of funding was distributed, these still result in small overall estimated price impacts (point estimates of less than 1 percent).

Taken together, these results help mitigate concerns that benefits to recipients could be driven by negative spillovers to non-recipient households, at least on the domains (subjective well-being, assets, expenditure and prices) examined here. In the near term, ongoing work will estimate effects on a wider set of household outcomes (including all those pre-specified as part of Haushofer et al. (2017) using this follow-up data. In addition, future work will study longer-term effects 4-5 years after the cash transfer distribution.

A Details of variable construction

A.1 Subjective Well-being

Summary measure: weighted, standardized index (following Anderson (2008)) separately for eligible and ineligible households, of the following measures:

- Depression (CES-D): we used a 10-question scale at endline and a 20-question scale at endline. Greater values correspond to increased depression. We negatively code this outcome when combining into the subjective well-being index
- Happiness (WVS): greater values indicate greater happiness
- Life Satisfaction (WVS): greater values indicate greater life satisfaction
- Perceived Stress Scale (Cohen's): greater values indicate greater stress; we negatively code for inclusion in the index

A.2 Assets

- Total non-land, non-home assets, net of lending: reported value of household's transportation (bicycles, motorcycles, car); furniture; TV, radio/casette player/CD player; agricultural tools; livestock; and other movable assets; plus household lending, net of household borrowing
- Total home value: this comes from respondent answers to a question on how much they would expect it to cost to build a home like theirs. As land and housing markets are quite thin in this area, generating reliable estimates of housing costs is a challenge.

A.3 Expenditure

- Total expenditure: Annualized sum of food expenditure in the last 7 days, frequent non-food purchases in the last month, and infrequent purchases in the last 12 months.
- Food expenditure in the last seven days: amounts on cereals, roots, pulses, vegetables, meat, fish, dairy/eggs, sugars, soft drinks, food outside the home, other food

- Frequent non-food purchases in the last month: airtime and internet; transport, clothing, recreation / entertainment; personal and household care items; and fuel, electricity and water.

- Infrequent purchases in the last 12 months: spending on housing (rent/mortgages, home maintenance and construction), religious events and charitable donations, social expenses (including weddings and funerals), medical expenses, schooling expenses, and spending on household durables.

- Temptation goods: alcohol and tobacco spending in the last 7 days; gambling expenditure in the last month

References

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.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics* 92, no. 1 (September): 1–45.

———. 2008. "Spatial Econometrics." In *The New Palgrave Dictionary of Economics,* Second Edition, edited by Steven N. Durlauf and Lawrence E. Blume, 7:741–747. Houndsmills: Palgrave Macmillan.

Evans, David K., and Anna Popova. 2014. "Cash Transfers and Temptation Goods: A Review of Global Evidence." World Bank Policy Research Working Paper 6886, May.

Haushofer, Johannes, Edward Miguel, Paul Niehaus, and Michael Walker. 2016. "Pre-analysis Plan for Midline Data: General Equilibrium Effects of Cash Transfers." May.

———. 2017. "GE Effects of Cash Transfers: Pre-analysis plan for household welfare analysis." July.

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics* 131 (4): 1973–2042. ISSN: 0033-5533, 1531-4650.

Jack, William, and Tavneet Suri. 2011. "Mobile Money: The Economics of M-PESA." NBER Working Paper No. 16721, January.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *The Review of Economics and Statistics* 91 (3): 437–456.

Mbiti, Isaac, and David N. Weil. 2015. "Mobile Banking: The Impact of M-Pesa in Kenya." In *African Successes, Volume III: Modernization and Development,* 247–293. NBER Chapters. National Bureau of Economic Research, Inc, March.

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

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217.