General equilibrium effects of cash transfers: experimental evidence from Kenya

Dennis Egger Johannes Haushofer Edward Miguel
Paul Niehaus Michael Walker*

December 22, 2020

Abstract

How large economic stimuli generate individual and aggregate responses is a central question in economics, but has not been studied experimentally. We provided one-time cash transfers of about USD 1000 to over 10,500 poor households across 653 randomized villages in rural Kenya. The implied fiscal shock was over 15 percent of local GDP. We find large impacts on consumption and assets for recipients. Importantly, we document large positive spillovers on non-recipient households and firms, and minimal price inflation. We estimate a local transfer multiplier of 2.4. We interpret welfare implications through the lens of a simple household optimization framework.

*Egger: University of California, Berkeley; Haushofer: Princeton University, NBER, Busara Center for Behavioral Economics, and Max Planck Institute for Collective Goods; Miguel: University of California, Berkeley, NBER, and CEGA; Niehaus: University of California, San Diego, NBER, and CEGA; Walker: University of California, Berkeley, and CEGA. Niehaus is a co-founder, former president (2012 - 2017) and chairman of the board of GiveDirectly. Acknowledgements: We thank the study participants for generously giving their time; Justin Abraham, Aakash Bhalothia, Christina Brown, Genevieve Deneoux, Tilman Graff, Max Lauletta, Michelle Layvant, Layna Lowe, Anya Marchenko, Gwyneth Miner, Max Mueller, Priscilla de Oliviera, Robert On, Rachel Pizatella-Haswell, Emaan Siddique, Zenan Wang, Francis Wong and Kejian Zhao for excellent research assistance; Innovations for Poverty Action for data collection fieldwork; the Busara Center for Behavioral Economics for piloting work; GiveDirectly for fruitful collaboration; and Martin Aragoneses, Vittorio Bassi, Abhijit Banerjee, Chris Blattman, Giacomo DiGiorgi, Esther Duflo, Yuriy Gorodnichenko, Seema Jayachandran, Adriana Lleras-Muney, Atif Mian, Emi Nakamura, Ben Olken, Jon Steinsson, and numerous conference and seminar participants for comments. This research was supported by grants from the National Science Foundation, International Growth Centre, CEPR/Private Enterprise Development in Low-Income Countries (PEDL), the Weiss Family Foundation, and an anonymous donor. The GiveDirectly program was supported by The Pershing Square Foundation. Walker gratefully acknowledges financial support from the NSF Graduate Research Fellowship (Grant No. DGE 1106400). The study received IRB approval from Maseno University and the Univ. of California, Berkeley. AEA Trial Registry RCT ID: AEARCTR-0000505.

1 Introduction

Tracing out the pattern of transactions in an integrated economy, and their contribution to aggregates of interest like overall output or well-being, has long been a central task of economic analysis. For instance, there has been interest in understanding the aggregate impacts of fiscal stimulus and cash infusions for decades (Keynes 1936), and a growing body of empirical evidence from rich countries shows that fiscal multipliers can sometimes be positive and large, based on non-experimental variation generated by policy changes (Chodorow-Reich 2019; Nakamura and Steinsson 2014; Suarez Serrato and Wingender 2016; Auerbach, Gorodnichenko, and Murphy 2019). Until now, however, these issues have not been subjected to experimental examination.

There is also renewed interest in related topics in development economics with the rise of large-scale cash transfer programs, which have now been implemented in scores of low and middle income countries.[1] A large literature on the impacts of these transfers has developed, employing well-identified experimental and quasi-experimental designs. These studies have documented effects on a broad range of behavioral responses among treated households, including consumption, earnings, assets, food security, child growth and schooling, self-reported health, female empowerment, and psychological well-being (Haushofer and Shapiro (2016), Baird, McIntosh, and Ozler (2011), and Bastagli et al. (2016)). Yet there is limited evidence on the aggregate economic impacts or welfare consequences of such policies (for exceptions, see Angelucci and De Giorgi (2009), Cunha, De Giorgi, and Jayachandran (2018), and Filmer et al. (2018)).

The present study was prospectively designed to unite these two disparate literatures by experimentally studying the aggregate impacts of large cash stimulus programs. We designed and carried out a large-scale experiment in rural Kenya that provided one-time cash transfers worth roughly USD 1000 (distributed by the NGO GiveDirectly) to over 10,500 poor households in a sample of 653 villages with a population of roughly 300,000. The implied fiscal shock was large, as the cash transfers amounted to over 15% of GDP in treatment villages during the peak 12 months of the program.

Beyond its fiscal scale, at least three aspects of the project represent advances on most existing work. First, we generated substantial spatial variation in the intensity of transfers by deliberately randomizing the allocation of cash transfers not just across households or villages (as is typical), but also across geographic sublocations (groups of 10–15 villages), thereby increasing our ability to detect aggregate impacts. Second, we carried out unusually

1. 97% of developing countries in Europe, Latin America and Sub-Saharan Africa have some type of cash transfer program (World Bank 2017).

extensive original data collection, giving us greater visibility into the chain of causal effects linking cash transfers to aggregate outcomes in a complex and interconnected economy. Our household and enterprise censuses of the study area count 65,385 households and 12,095 non-farm enterprises. Within this unusually large sampling frame, we gathered detailed panel (longitudinal) data on household receipt of the transfer; household consumption expenditure patterns (representative for both recipient and non-recipient households); local enterprise production, employment and revenue; labor market conditions; as well as especially high-frequency (monthly) and spatially disaggregated market data on prices. Third, we interpret the results through the lens of a theoretical framework that highlights the links between the individual empirical results, the aggregate transfer multiplier, and welfare in this setting.

Following earlier studies, we first document large direct impacts on households that received transfers, including increases in consumption expenditures and holdings of durable assets eighteen months after the start of transfers. We do not observe meaningful changes in total labor supply among recipient households.

Enterprises in areas that receive more cash transfers also experience meaningful revenue gains, in line with the increases in household expenditure. Interestingly, sales increased without noticeable changes in firm investment behavior (beyond a modest increase in inventories), and did not increase differentially for firms owned by cash recipient households relative to non-recipients. Both patterns suggest a demand- rather than investment-led expansion in economic activity. Increased enterprise revenue in turn translates into moderate increases in wage bills and profits. Methodologically, one important feature of the enterprise (and to a lesser extent household) results is that they are largely driven by the overall intensity of treatment in nearby communities, not solely by the treatment status of the village in which the enterprise is located. This suggests that common study designs which aim to identify spillover effects by clustering treatment at the village level and assuming no spillovers across villages may be mis-specified, at least in densely populated areas such as the one we study.[2]

Despite not receiving transfers, non-recipient households also exhibit large consumption expenditure gains: their annualized consumption expenditure is 13% higher eighteen months after transfers began, an increase roughly comparable to the gains contemporaneously experienced by the recipient households.[3] Increased spending is not financed by dissaving, but more likely results in part from the income gains experienced by local firms' owners and workers. Indeed, non-recipients' income gain is driven largely by increases in wage labor earnings, consistent with the fact that enterprise wage bills increase. In a reassuring check,

2. For example, households are located within 2 km of seven other villages on average.

3. We note below that consumption gains among recipient households are likely to have been larger in the period immediately following transfer receipt.

the magnitude of per capita consumption gains among local households lines up roughly with the per capita revenue gains among local firms. On some level this is unsurprising, as increases in local consumption expenditures were spent somewhere; our contribution is to carefully document how such spending spreads locally through a low-income economy.

A further issue is the extent to which transfers affect local prices (as for example Cunha, De Giorgi, and Jayachandran (2018) show in Mexico), and thus the extent to which the effects described above are nominal or real. We study this question through careful monthly measurement of prices for scores of local commodities, consumer goods, and durable goods, as well as prices for inputs like labor and capital. For inputs, we find positive point estimates, but they are economically moderate in magnitude and not always statistically significant. For outputs, we document statistically significant, but economically minimal, local price inflation. Average price inflation is 0.1%, and even during periods with the largest transfers, estimated price effects are less than 1% and precisely estimated across all categories of goods.

We next ask what these effects imply for the aggregate level of economic activity, computing a local transfer multiplier. A standard macroeconomic framework would predict that large multipliers are possible in our rural Kenyan setting: it is a largely closed local economy within a currency union receiving external transfers, with incomplete markets and a large share of hand-to-mouth consumers (Farhi and Werning 2016). Using an expenditure-based approach that takes advantage of our data on the consumption expenditures of representative samples of both recipient and non-recipient households as well as investment by local firms, we estimate a local transfer multiplier of 2.5. A dual income-based approach, relying on distinct and complementary measures of labor and capital income, enterprise profits, and taxes, yields a similar estimate of 2.3. These estimates are broadly in line with what a simple model would imply from households' marginal propensity to consume local value added, which we estimate to be approximately 0.76 over the study period. These results contribute to an active recent empirical literature that estimates multipliers.[4]

A core contribution of this study is thus to exploit a randomized experiment to estimate an important macroeconomic quantity (see Muralidharan and Niehaus (2017) for a related discussion). A notable aspect of our approach is the fact that transfers came from donors out-

4. Our estimates are somewhat larger than those from a structural simulation, which predicted that the local multiplier from cash transfers in rural Kenya could range from 1.6 to 1.9 (Thome et al. 2013), and are similar in magnitude to non-experimental estimates from a cash transfer program in Mexico (1.5 to 2.6) (Sadoulet, Janvry, and Davis 2001). They are also somewhat larger than recent estimates of the fiscal spending multiplier (which is distinct from the transfer multiplier, since households can save part of the transfer) derived from cross-sectional US policy variation, which often range from 1.5-2.0 (Chodorow-Reich 2019), and from Brazil, which are close to 2 (Corbi, Papaioannou, and Surico 2019). Pennings (2020) focuses on the US transfer multiplier. Kraay (2014) estimates fiscal multiplier estimates less than one when donor lending is used as an instrument for national government spending in developing countries.

side the study area, rather than being internally tax- or debt-financed; the latter is typically the case in the US programs studied, and may complicate the interpretation of consumption responses due to contemporaneous tax incidence or Ricardian equivalence issues. The targeting of the transfers to just one region within the larger Kenyan economy also allows us to abstract away from monetary policy and exchange rate responses, simplifying analysis relative to the study of national stimulus policies. A limitation of our approach is that we observe partial data in the months immediately following the transfers, which reduces the precision of some estimates.

Few existing treatments of multipliers also explicitly examine their welfare implications.[5] We interpret the welfare implications of our results using a simple theoretical framework. Transfers directly increase the welfare of those who receive them by $1 per $1 received. General equilibrium effects impact welfare through two additional channels: changes in household budget sets (due for example to changes in wages, prices, or firm profits), and changes in peer behaviors that enter directly into own utility (due to externalities or public good provision). The value of budget set expansions depends on what drives them: expansions due to increases in productivity are worth $1 per $1, while expansions due to increased factor supply (e.g., labor) come at a partially offsetting opportunity cost. Interpreted through this lens, the results generally suggest that non-recipients were made better off by an expansion in their budget sets driven largely by increased factor productivity, as opposed to factor supply. Externality effects are positive or null, with one possible exception: positive spillovers were large enough that village-level asset inequality increased slightly, which may affect well-being if households have preferences over their relative socioeconomic standing.

The constellation of findings raises an intriguing question about how the economy absorbed such a large shock to aggregate demand, and in particular how it did so without correspondingly large increases in the employment of land (which is in fixed supply), labor, or capital. One plausible, albeit speculative, possibility is that the utilization of these factors was "slack" in at least some enterprises (Lewis 1954). This seems plausible because in the retail and manufacturing sectors, where output responses were concentrated, the typical firm has a single employee (i.e., the proprietor), suggesting that integer constraints may often bind. In addition, many enterprises operate "on demand" in the sense that they produce only when they have customers, and the average non-agricultural enterprise sees just 1.9 customers per hour. In addition to retail, much manufacturing in this setting is on demand; for example, a mill owner waits for customers to bring their grain. The existence of slack may help account for the large multiplier we estimate, as has also recently been argued in US data (Michaillat and Saez 2015; Murphy 2017).

5. Recent and notable exceptions include Mankiw and Weinzierl (2011) and Sims and Wolff (2018).

2 Study design

2.1 Setting: rural western Kenya

The study took place in three contiguous subcounties of Siaya County, a rural area in western Kenya bordering Lake Victoria. The population in Siaya is predominantly Luo, the second largest ethnic group in Kenya, and while rural it is also relatively densely populated, with 395 people per km^2 compared to a Kenyan average of 91. The main national road running from the port of Mombasa to Nairobi and on to Kampala passes through the study area, likely helping to integrate it into the regional economy.

The NGO GiveDirectly (GD) selected the study area based on its high poverty levels. Within this area GD selected rural (i.e., not peri-urban) villages in which it had not previously worked.[6] This yielded a final sample of 653 villages spread across 84 sublocations (the administrative unit above a village). The mean village consists of 100 households, and at baseline, the average household had 4.3 members, of which 2.3 were children. The average survey respondent was 48 years old and had about 6 years of schooling. 97% of households were engaged in agriculture; at endline, 49% of households in control villages were also engaged in wage work and 48% in self-employment.

Transfers and data collection took place from mid-2014 to early 2017, a period of steady economic growth, relative prosperity, and political stability in Kenya. The World Bank reports annual per capita GDP growth rates ranged between 2.4 to 3.4 percent. All data collection concluded months prior to the August 2017 national election.

2.2 Intervention: The GiveDirectly (GD) Cash Transfer Program

GD provides unconditional cash transfers to poor households in low-income countries. For the purpose of this study, to be eligible for transfers, households had to live in homes with thatched roofs, a simple means-test for poverty. In treatment villages, GD enrolled all households that met this criterion ("eligible" households) as classified by their field staff through a village census, and confirmed via two additional visits (see Appendix B.1).

Approximately one-third of all households were eligible. These households received a series of three transfers totaling KES 87,000, or USD 1,871 PPP (USD 1,000 nominal), via the mobile money system M-Pesa, which is widely used in Kenya. (Registering for M-Pesa was a prerequisite for receiving transfers; households without a mobile phone were given the option to purchase one from GD staff with the cost deducted from their transfer.) Households selected the member they wished to receive the transfers. The total transfer

6. The listing of villages was based on the 2009 National Population Census; enumeration areas (which typically correspond to a single village) were treated as villages by GD and this study.

is large, corresponding to 75 percent of mean annual household expenditure in recipient households. In aggregate, the transfers were equivalent to approximately 17 percent of annual GDP (based on our data described below) in the treated areas during the peak 12 months of disbursements, and to 24 percent of annual GDP during the full 24 month rollout period. Although small in relation to overall Kenyan GDP in 2015 (<0.1%), locally this is thus a larger relative shock than most government transfer programs, e.g., the ARRA programs studied in the recent US fiscal multiplier literature, see Chodorow-Reich (2019).

Transfers were made in a series of three payments as follows: a token transfer of KES 7,000 (USD 151 PPP) was sent once a majority of eligible households within the village had completed the enrollment process, followed two months later by the first large installment of KES 40,000 (USD 860 PPP). Six months later (and eight months after the token transfer), the second and final large installment of KES 40,000 was sent. The median treated household received its token transfer 47 days after being registered for the program; transfers should be interpreted as anticipated during that period to the extent recipients believed GD's promises.[7] The transfers were non-recurring, i.e., no additional financial assistance was provided to recipient households after their final installment, and they were informed of this up front. Households in control villages did not receive transfers.

2.3 Experimental design

To identify spillovers both within and across villages, we employed a two-level randomization design (Figure A.1, Panel A). First, we randomly assigned sublocations (or in some cases, groups of sublocations) to high or low saturation status, resulting in 33 high- and 35 low-saturation groups. Within high (low) saturation groups we then randomly assigned two-thirds (one-third) of villages to treatment status. We also randomized the order in which treatment was rolled out to treated villages. Within treatment villages, all eligible households received a transfer.[8] This design induces variation in treatment intensity across space due both to the variation in sublocation treatment intensity, and random variation in the location of treated villages within sublocations. Figure A.2 illustrates that there is considerable

7. The precise timing of the first transfer was uncertain, and we believe that recipients may also have perceived the first transfer as less certain than the subsequent ones. However, we do not know of any borrowing against future GD payments, and credit markets are imperfect in our context. In the earlier GiveDirectly study, Haushofer and Shapiro (2016) find evidence for both savings and credit constraints: households that received lump-sum transfers were more likely to own large durable assets at endline than households receiving monthly transfers, even though the total transfer amounts were the same. This suggests that households had trouble borrowing against the promise of the future transfer and saving the early installments. This is also consistent with US evidence finding no anticipation in advance of the receipt of economic stimulus payments (Broda and Parker 2014). We therefore consider all transfers symmetrically in our dynamic regressions and leave analysis of potentially differing effects across installments for future work.

8. Full details of the randomization are in Appendix B.2.

variation both across and within sublocations in the share of neighboring villages treated.[9]

3 Data and empirical specifications

We conducted four types of surveys, of households, enterprises, market prices, and local public goods. Results from the public goods surveys are presented primarily in a separate paper (Walker 2018), and discussed briefly here. We filed several pre-analysis plans for this project; for details on the PAPs and where we go beyond these plans, see Appendix I.

3.1 Household data

We first conducted a baseline household census in all villages, which serves as a sampling frame and classifies household eligibility status. The census was designed to mimic GD's censusing procedure but was conducted by independent (non-GD) enumerators across both treatment and control villages for consistency. Throughout this paper, we base our analysis on village membership, household definitions and eligibility as classified by our project data collection field staff. In all, the census identified 65,385 households with a total baseline population of 280,000 people in study villages.

We conducted baseline household surveys within one to two months after the census and before the distribution of any transfers to a village (Figure A.1, Panel B).[10] We used census information to sample at random eight eligible and four ineligible households per village to survey. When households contain a married or cohabiting couple, we randomly selected one of the partners as the target survey respondent. Due to time and budget constraints, we sought to complete all baseline household surveys in a single day. If a household on our sampling list was not available on that day, we instead surveyed a randomly-selected replacement household with the same eligibility status. We conducted a total of 7,848 baseline household surveys between September 2014 and August 2015.[11] The survey contained detailed modules on economic activities, asset ownership, psychological well-being, health and nutrition. A large array of baseline characteristics are balanced across treatment and control villages (Table C.2, column 2).

Endline household survey data was collected between 9 and 31 months after each household's "experimental start date," meaning the month in which transfers were expected to

9. Figure B.1 provides a higher-resolution example for two villages.

10. In a few cases, baseline surveys were conducted before the distribution of transfers but after GD had held meetings in the village informing households that it would be a treatment village.

11. Of this total, 6,510 households were on the initial sampling list, and 1,338 were randomly-selected replacement households.

start in a village assigned to treatment, regardless of their actual treatment timing.[12] The 5th/95th percentiles of timing ranged from 12 to 27 months, and the median survey was conducted 19 months after the experimental start month, or about 11 months after the distribution of the last lump sum transfer (Figure A.1, Panel B). This timing implies that some but relatively few households were surveyed in the months immediately following cash transfer receipt, an issue we return to below when estimating the transfer multiplier.

Endline household surveys targeted all households on the initial sampling lists (including those missed at baseline), along with replacement households that were surveyed at baseline. For households that had been surveyed at baseline, we attempted to survey the individual who was the baseline respondent. We conducted a total of 8,242 endline household surveys between May 2016 and June 2017.[13] We achieved high tracking rates at endline, reaching over 90% of eligible and ineligible households in both treatment and control villages, and these rates do not systematically vary by treatment status (Table C.1). The only subgroup difference of note is that we are slightly less likely to find ineligible households that were initially surveyed at baseline in high saturation sublocations (see Appendix C.2 for more information). In addition to the baseline modules, endline surveys collected more detailed data on household expenditures and crop production, additional psychological scales (in particular, related to stress and hope), and female respondents surveyed by a female enumerator were also administered a module on female empowerment and gender-based violence.

3.2 Empirics: recipient households

If the general equilibrium effects of transfers were fully contained within administrative units (here, villages and sublocations), then an appropriate specification would be

$$y_{ivs} = \alpha_1 Treat_v + \alpha_2 HighSat_s + \delta_1 y_{ivs,t=0} + \delta_2 M_{ivs} + \varepsilon_{ivs}, \tag{1}$$

where y_{ivs} is an outcome of interest for household i in village v in sublocation s.[14] $Treat_v$ is an indicator for residing in a treatment village at baseline, and $HighSat_s$ an indicator for being in a high-saturation sublocation. Here α_1 captures the total average treatment

12. The order in which villages were visited by GD and the research team was randomized within subcounties. We calculate the start and end months of when GD started transfers to villages within a subcounty, and then, across these months, evenly assign both treatment and control villages experimental start months based on the random ordering.

13. This includes 7,019 initially sampled and 1,223 replacement households. Of the initially sampled households, 1,015 had been missed at baseline. The main analysis focuses on the "initially sampled" (which includes those missed at baseline) and "replacement" households; results are similar using only originally sampled households (available upon request).

14. When we examine individual-level outcomes using Equation (1), we define treatment status and eligibility on the basis of the household in which the individual lives.

effect for households in treatment versus control villages, including both the direct effect of treatment (for eligible households) and any within-village spillovers; note that our design does not allow us to identify these separately. α_2 is a relatively coarse way to assess cross-village spillovers, as it does not utilize all experimental variation. We include the baseline value of the outcome variable ($y_{ivs,t=0}$), when available, to improve statistical precision.[15] We cluster standard errors at the village level, and weight observations by inverse sampling probabilities to be representative of the population of eligible households.

Overall, we view Equation 1 as a useful benchmark but unlikely to capture well the spatial variation in treatment intensity evident in Figure A.2. This is because in our study area villages are relatively close to each other; sublocation boundaries are not "hard" in any sense nor reflective of salient ethnic or social divides; and because our data indicate that there is extensive economic interaction in nearby markets regardless of sublocation. To better capture spillovers, we therefore estimate models in which a household's outcomes depend on the amount of money distributed in its own and other geographically proximate villages:

$$y_{iv} = \alpha + \beta Amt_v + \sum_{r=2}^{R} \beta_r Amt_{v,r}^{\neg v} + \delta_1 y_{iv,t=0} + \delta_2 M_{iv} + \varepsilon_{iv}. \tag{2}$$

The novel terms here are the amount Amt_v of cash per capita transferred to one's own village v over the entire study, and the amount $Amt_{v,r}^{\neg v}$ of cash per capita transferred to villages other than v in a series of bands with inner radius $r - 2$ km and outer radius r km around the village centroid. We normalize both to be measured as a share of per capita GDP.[16] The Amt variables depend on both the random assignment of villages to treatment and also on the endogenous share of households in those villages eligible for transfers, so we instrument for them using the own-village treatment indicator $Treat_v$ and the share $s_{\neg v,r}^{e,t}$ of eligible households in each band assigned to treatment. To account for spatial correlation, we calculate standard errors using a uniform kernel up to 10 km (Conley 2008).[17]

Because we had no a priori knowledge of the relevant distances over which general equilibrium effects might operate, we pre-specified an approach in which we first estimate a series of nested models varying the outer limit R of the spatial bands from 2 km to 20 km in

15. For observations where the baseline value is missing, we include an indicator variable equal to one denoting a missing value (M_{ivs}), and set the baseline value of the outcome variable equal to its mean.

16. We use an expenditure-based measure of GDP that is described in Section 5, which we convert to a per-capita measure based on household census data from our study area, and augmented with data from the GiveDirectly census and the 2009 Kenya National Population Census when necessary. Per capita GDP in low saturation control villages is 637 USD PPP (2727 USD PPP per household); see Appendix B.3.

17. We also conduct Fisher randomization tests for all specifications, where we re-randomize cash transfers across sublocations and villages as well as their roll-out over time as in our experiment and test against the sharp null that effects are zero. Conclusions are robust to this alternative method of inference.

steps of 2 km, and then select the one which minimized the Schwarz Bayesian Information Criterion (BIC). We report estimates of Equation (2) using the selected outer limit $\bar{R}$. As it turns out, this algorithm selects only the innermost 0–2 km band for almost all outcomes.[18]

Equation (2) correctly identifies the overall effects of the intervention if transfers delivered outside the radius $\bar{R}$ have no effect on i. If not – if, for example, all households were affected to some extent by all transfers in the study area – then the estimated effects are *relative* to these "ambient" effects. The BIC selection procedure determines how reasonable this identifying assumption is by omitting ring $R + 2$ from the model if it has little explanatory power for the outcome. In principle, this could be either because variation in treatment intensity at that distance has a precisely estimated but small effect on the outcome (in which case it is a good "control"), or because there simply is not much variation in treatment intensity at that distance (in which case we cannot be sure). Given this, it is important to note that our design generates substantial variation in treatment intensity even at larger distances. Transfers in the 2 km buffer, which the BIC always includes, range from 0 to 27% of GDP, with a 10-90 percentile range of [4%, 15%]. Even in the 4 to 6 km buffer, which the BIC *never* selects for any primary outcome, the 10-90 percentile range remains wide, at [3%, 10%]. This suggests that a subset of our villages can reasonably serve as "pure controls."

We also examine in Appendix F how sensitive our main conclusions are to fixing larger maximal radii $\bar{R}$ than the BIC selects, which implies a more conservative definition of the "control group." While there is some variation from outcome to outcome, overall the effects are quite stable as we add the 2-4 km band and fairly stable as we add the 4-6 km band, though as expected standard errors often become much larger. We generally cannot reject that these estimates are different statistically from those estimated using the BIC-optimal bandwidth, giving us greater confidence in the latter. Finally, note that we typically estimate spillovers of the same 'sign' as the direct effects, which suggests that any remaining bias in our estimates likely leads us to understate, rather than overstate, overall effects. All told, we view the problem of estimating spatial effects as unlikely to admit a perfect solution, but believe that our study design and econometric specification allow us to advance meaningfully relative to most existing work.

18. Note that this model selection step introduces some circularity, as we first determine the distances at which effects occur, and then estimate effects at those distances. We check that inference is robust to this model selection, and to alternative approaches more generally. First, we calculate exact p-values using a Fisher permutation test (i.e. randomization inference, see Appendix F.4). Second, we conduct repeated 50-50 splits of the data into training and test sets, using the training data from each split to perform the BIC step and the test set to estimate parameters, and record the proportion of times that the resulting estimates lie within the 95% confidence intervals we report here (Appendix F.2). Third, we estimate effects holding the spatial radius fixed at 2 km, 4 km, and 6 km, respectively, thus eliminating the model selection step (Appendix F.1). Taken together, these results provide reassurance that our methods yield valid inferences.

We estimate Equation (2) for all *eligible* households and then use it to obtain estimates of the total effect on *recipient* households, which we report as "Recipient Households" in tables. By "total effect" we mean how the households' outcomes differ from what they would have been in the absence of the intervention. We calculate these by multiplying the estimated coefficients from Equation (2) by the average values of the regressors, i.e., $\hat{\beta} \cdot \left(\overline{Amt_v} | i \text{ is an eligible recipient}\right) + \sum_{r=2}^{\bar{R}} \hat{\beta}_r \cdot \left(\overline{Amt_{v,r}^{\neg v}} | i \text{ is an eligible recipient}\right)$ for all radii bands up to the selected $\bar{R}$. This effect allows for across-village spillovers in addition to direct effects and within-village spillovers.[19] As a benchmark, we also report estimates of α_1 from Equation (1), which is the total treatment effect if all spillovers are contained within villages (a common identifying assumption).

3.3 Empirics: non-recipient households

We use an analogous approach to estimate total effects on non-recipient households, which include both eligible households in control villages and ineligible households in all villages. Specifically, we estimate

$$y_{iv} = \alpha + \sum_{r=2}^{\bar{R}} \beta_r^1 Amt_{v,r} + \sum_{r=2}^{\bar{R}} \beta_r^2 (Amt_{v,r} \cdot Elig_{iv}) + \gamma Elig_{iv} + y_{iv,t=0} \cdot \delta + \varepsilon_{iv}. \tag{3}$$

This specification modifies Equation (2) as follows. First, because non-recipient households do not experience direct effects, we no longer separate own-village effects and across-village spillovers: we drop Amt_v and replace $Amt_{v,r}^{\neg v}$ with $Amt_{v,r}$, so that spillovers work entirely through β_r^1 and β_r^2. Second, we include an indicator for eligibility status and its interaction term with amounts to allow for spillovers to differ by eligibility status (recall that eligible households in control villages are non-recipients). As above, we instrument for $Amt_{v,r}$ using the share of eligible households assigned to treatment within the corresponding band. When available, we include the baseline value of the outcome variable. We report the average total effect on non-recipients as a population-weighted average of effects for the two groups.[20]

19. Appendix C.5 provides an example of this for outcomes in Table 1. We also consider the possibility that effects are non-linear in the per-capita amounts transferred. Figure A.3 presents non-linear estimates of equation 2 for two key outcomes, total consumption and firm revenue. The relationships appear roughly linear, and we cannot formally reject linearity. We conduct the same test for our 10 pre-specified primary outcomes and eligible / ineligible households separately, and cannot reject linearity at the 10% significance level for any of them.

20. This is calculated as $s^{e,c}\left(\sum_{r=2}^{\bar{R}} \left(\hat{\beta}_r^1 + \hat{\beta}_r^2\right) * \left(\overline{Amt_{v,r}} | i \text{ is an eligible non-recipient}\right)\right) + s^i \left(\sum_{r=2}^{\bar{R}} \hat{\beta}_r^1 * \left(\overline{Amt_{v,r}} | i \text{ is ineligible}\right)\right)$, where $s^{e,c} = 1 - s^i$ is the population share of eligible non-recipient village households among all non-recipient households, and the $\hat{\beta}_r^1$ and $\hat{\beta}_r^2$ terms come from Equation (3).

3.4 Enterprise data

We employ several complementary sources of data on enterprises. First, we use detailed agricultural and self-employment modules from the household surveys. The agriculture module covers crop-by-crop agricultural production, sales, employment, and input costs; the self-employment module covers revenues, profits, hours worked, and some costs and investments for enterprises run by household members. These data are representative of enterprises owned locally (i.e., by residents of the study area) and allow us to clearly attribute profits to their residual claimants. They do not capture enterprises owned by people living outside the study area, which we capture separately through the enterprise census and surveys.

Specifically, we conducted censuses and surveyed a representative subset of all non-farm enterprises at baseline and endline (see Appendix D for details). The endline census was conducted between November 2016 and April 2017, covering both enterprises identified at baseline and newly established enterprises. This served as the endline survey sampling frame; we randomly sampled up to 5 enterprises per village, stratified by those operating from within and outside of homesteads. Surveys covered revenue, profits, employees, wages, some other costs, and taxes paid. The main endline sample includes 1,699 enterprises operated from within and 1,442 from outside the homestead (both from enterprise surveys), as well as 7,899 agricultural enterprises from the household survey. Enterprise characteristics appear balanced across treatment and control villages at baseline (Table D.3): there are no statistically significant differences at the 5% level, and two of 18 are significant at 10%, which we would expect by chance.

This integrated approach to household and enterprise surveying allows us to match firms to their owners. We match all agricultural enterprises (as found via household surveys), and 56% of non-agricultural enterprises, for a total of 93% of all enterprises. Based on this match, we estimate that enterprise activity is highly localized, with 92% of total profits and 87% of revenues accruing to owners who live within the village in which the enterprise operates.

3.5 Empirics: enterprises

We estimate enterprise-level effects using versions of Equations (1) and (2), with radii bands selected as above, interacting right hand side variables with enterprise type (Appendix D.2 lists the full specifications). We include village-level means rather than enterprise-level values of the lagged dependent variable given that the enterprise surveys were repeated cross-sections. We carry out estimation using inverse probability weighting, accounting for enterprise type, except in some cases where we also revenue-weight outcomes. As above, we calculate and report average total effects, weighting effects for the three enterprise types, namely, agricultural enterprises, non-farm enterprises operating within homesteads, and non-farm enterprises operating outside the home; we typically pool data across all enterprise types, except when we do not observe some outcomes for agricultural enterprises. To facilitate comparisons between the household and enterprise results, we also normalize effects as per-household rather than per-enterprise.[21] To examine extensive margin effects, we estimate village-level analogues to this approach with the total number of enterprises censused (per household) as the dependent variable.

3.6 Price data

We measure consumer goods prices using monthly surveys of commodity prices in local markets. These surveys were conducted over the course of 2 to 2.5 years in all 61 markets in the study area (and neighboring towns) with at least a weekly market day, for a total of 1,586 market-by-month observations and 311,138 non-missing price observations. We have market price data prior to the disbursement of any local cash transfers for all markets, providing an appropriate baseline for the panel data econometric analysis detailed below, and allowing for the inclusion of market fixed effects. These include market centers located in towns, and so will appropriately reflect the impacts of households (potentially) traveling to towns to spend their transfers. Figure A.2 shows the substantial variation in treatment intensity around markets, as well as the heterogeneity in village proximity to markets. The average village had 0.7 markets located within 2 km and 2.3 markets within 4 km, again indicating the rather high density of residential settlement. Household respondents report an average commuting time to their preferred market of 33 minutes, where over 75% walk to the market.[22]

Market surveys collected prices for 70 relatively homogeneous products, including food (grains, vegetables, fruit, meat), livestock (goats, sheep), hardware (nails, paint), "duka" kiosk store products (non-food and packaged food), and others (e.g., fuel, health items, household items, and farming implements). We collected quotes from three vendors of each product in each market in each month, and use the median for each product-market-month. We then calculate linear log-price indices by weighting prices by household expenditure

21. Specifically, we calculate $\frac{1}{n_{hh}} \sum_g \widehat{\Delta y}_e^g * n_{ent}^g$, where n_{hh} is the total number of households across all control villages (column 2) or treated villages (column 3), $\widehat{\Delta y}_e^g$ is the estimated average effect ($\hat{\beta} * \bar{X}$) for enterprise type g, and n_{ent}^g is the number of enterprises of type g in the control or treated villages.

22. Enterprises in markets account for 65% of non-agricultural enterprise revenue, based on a 2019 census of enterprises. We did not collect price data as comprehensively from the minority of enterprises located outside of markets and dispersed within villages, both for logistical reasons and because their products tend to be less standardized. That said, estimated impacts on the prices of two common services these enterprises offer, tailoring and maize grinding, are if anything smaller than estimated effects on our main market price index (Table E.5).

shares.[23] We also examine effects on component subcategories of goods, which include: food items; non-food non-durables (such as soap, cooking fat, and firewood); durables (such as iron sheets and jerry cans used for transporting water or fuel); livestock; and temptation goods.[24]

We measure prices of the major factors of production using household survey data on wages, land prices, and interest rates on formal and informal borrowing and lending. Because compositional changes in these inputs may be important, we examine quantity and price effects side by side.

3.7 Empirics: prices

We estimate effects on consumer goods prices using both spatial and temporal variation in the amount of cash distributed around each market. In contrast to the household and enterprise data, our repeated measurement of prices, both before and after the start of cash distributions, allows us to estimate equations that include market fixed effects; in Appendix E.2.3, we demonstrate robustness to using an analytical approach analogous to that used with household and enterprise data. Specifically, we estimate

$$p_{mt} = \sum_{r=2}^{\bar{R}} \sum_{l=0}^{M} \beta_{rl} Amt_{m(t-l),r} + \alpha_m + \lambda_t + \varepsilon_{mt} \quad (4)$$

where p_{mt} is a price outcome for market m in month t. $Amt_{m(t-l),r}$ is the per-household amount transferred within band $r-2$ to r km around market m in month $t-l$, expressed as a fraction of GDP. We exploit our panel setup by conditioning on fixed effects for both markets (α_m) and months (λ_t). The former absorb any systematic price differences across markets as well as differences in the share of eligible households located around those markets, conditional on which treatment is randomly allocated. The latter account for seasonal differences and other time trends common to all markets. We again account for spatial correlation in calculating standard errors (Conley 2008). We determine both the relevant spatial distance R and the relevant temporal lag M over which price effects persist by minimizing an information criterion conceptually similar to that above, but adapted to account for the fact that the BIC cannot select between non-nested models (such as one with a high R and

23. We use expenditure data from the the Kenya Life Panel Survey (Baird et al. 2016) conducted in 2013-2014 in rural areas of Siaya and neighboring Busia county. We use the KLPS data because we did not collect a full expenditure module at baseline (due to project time and budget constraints) and prefer not to use endline expenditure data which are potentially endogenous. That said, results are nearly unchanged if we use consumption expenditure shares from non-recipient households in our endline survey.

24. The consumption expenditure measure of temptation goods includes alcohol, tobacco, and gambling. The price index includes the cost of cigarettes.

another with a high M). Specifically, and as pre-specified (Appendix I), we first select R while holding M fixed at 3 months by estimating models of the form

$$p_{mt} = \sum_{r=2}^{R} \beta_r \left(Amt_{mt,r} + Amt_{m(t-1),r} + Amt_{m(t-2),r} \right) + \alpha_m + \lambda_t + \varepsilon_{mt} \tag{5}$$

where R varies between 2 km and 20 km. We select the value $R = \bar{R}$ that minimizes the Schwarz BIC while imposing weak monotonicity. We then select the number of monthly lags M by estimating Equation (4) with $R = \bar{R}$ and choosing the model that minimizes the Schwarz BIC. This procedure selects only the 0-2 km band (and sometimes the 2-4 km band) around each market and a single temporal effect, implying that we only include contemporaneous transfers in estimating price effects. Appendix E.2 presents results for a specification where we impose $R = 4$ km and $M = 18$ months for robustness, and yields similar results.

Identification in Equation (4) comes from the roll-out of treatment across space *and* time, and the project's research design leads to substantial variation in both dimensions. As noted above, there is considerable variation in the total amounts of cash going to the 0–2 km ring around each market. Moreover, the multi-year nature of the market data covers periods both of intensive transfer distribution as well as times when no transfers were going out. As above, we are unable to capture price increases that radiate throughout the whole study area (compared to neighboring counties) over the entire period, but the highly localized nature of the price effects that we do detect suggests that any such effects are unlikely to be large.

We use estimates of Equation (4) to calculate two price effects. The implied *average treatment effect* (ATE) is the average price effect across all markets and all months in which any transfers went out to any market in the study area, i.e., during the study period of September 2014 to March 2017. This is simply equal to the estimated coefficients multiplied by the mean of the corresponding regressors of interest. The *average maximum transfer effect* is the average across markets of the estimated effect in the month in which the maximum amount of cash (as a share of GDP) was distributed into the selected radii bands (in other words, out to $\bar{R}$) from the market.

We focus on two sources of heterogeneous price effects. First, we classify goods into those that are more and less tradable, where the former include relatively easily transported, non-perishable items, and the latter include more difficult to transport or perishable items.[25]

25. For instance, more tradable goods include building materials (e.g., timber, cement, nails, iron sheets) and some household goods (soap, firewood, charcoal, batteries, washing powder), while less tradable goods include some food items (e.g., avocado, banana, cabbage, egg, pork, fish) and livestock. These classifications were undertaken based on feedback from Kenyan project staff, but there may, of course, be some ambiguity about specific items. The full pre-specified classification is in Appendix E.

Second, we classify markets into those with better or worse *market access*. Standard theory in international trade predicts that more integrated markets should be less likely to experience meaningful price changes following a local aggregate demand shock. We examine output price heterogeneity with respect to a commonly used metric of market access (MA_m)

$$MA_m = \sum_d \tau_{md}^{-\theta} N_d \approx \sum_{r=1}^{10} r^{-\theta} N_r \tag{6}$$

Geographic distance r is used to proxy for trade costs between origin market m and destination d, i.e., $\tau_{md} = r$. Destinations are 1 km radii bands around each market, with total population N_r in each buffer, and we follow Donaldson and Hornbeck (2016) in setting $\theta = 8$. Within quantiles of this metric, we calculate average and average maximum transfer treatment effects in the manner described above.[26]

We estimate effects on input prices using Equations (1), (2), and (3), as our input price data come from household surveys, and report the corresponding average treatment effects.

3.8 Empirics: dynamics

To capture the full dynamic response of outcomes and to calculate the multiplier, we estimate and then integrate effects on components over time.[27] For a flow variable x (e.g., consumption, investment, etc.), we first estimate the following specification, which is a dynamic extension of previous estimating equations:

$$x_{it,v} = \alpha_t + \sum_{s=0}^{9} \beta_s Amt_{v(t-s)} + \sum_{s=0}^{9} \gamma_s Amt_{v(t-s),\text{0-2km}}^{\neg v} + \varepsilon_{it,v}. \tag{7}$$

where $Amt_{v(t-s)}$ is the amount transferred to village v in quarter $t-s$, instrumented by a treatment indicator $Treat_v$ multiplied by the share of total transfers going to village v in quarter $t-s$, and analogously $Amt_{v(t-s),\text{0-2km}}^{\neg v}$ by the share $s_{v,\text{0-2km}}^{e,t,\neg v}$ of eligible households in the 0 to 2 km buffer around v (but not in village v) assigned to treatment multiplied by the share of transfers going to that group in quarter $t-s$. The coefficients in this model are identified by the fact that village treatment status was randomized, and the timing of both cash transfers and survey data collection was rolled out to villages in a randomized order. The main challenge is that the first household surveys started around 9 months after the experimental start date in each village, while enterprise surveys began after about 18

26. We also consider an alternative market access metric, namely, road access, defined as the inverse distance from the closest main road (as captured by Open Street Map), see Appendix E.2.2.

27. We stated our intention to estimate a multiplier in our pre-analysis plans, but did not fully specify the econometric approach for doing so.

months (see Figure A.1). With a few exceptions, recall periods are less than or equal to one month, so we often do not directly observe the initial response in flow variables in the months immediately after the first transfers went out, which is when we might expect to see some of the largest impacts on expenditure. Given that our specification treats each dollar transferred symmetrically, we can still estimate the local response during these early quarters because transfers to recipients rolled out over 8 months. Similarly, we estimate neighborhood effects using the substantial variation in the timing with which nearby villages were treated. However, we tend to obtain less precise estimates of responses in early quarters immediately following transfers, as they are estimated using less variation in treatment intensity compared to later quarters.

We then integrate dynamic effects on flow variables over time up to 29 months (10 quarters) after treatment. We compute the dynamic profile of treatment effects (or the impulse response function, IRF) using the coefficients estimated above and assuming that the treatment rolled out to recipient households as planned: the timing is a token transfer at time 0, a first lump-sum 2 months later, and a second 8 months after the token transfer. We compute this IRF separately for recipient and non-recipient households, and separately for the three categories of enterprise in both treatment and control villages. We then aggregate the quarterly estimates across all villages (using inverse population weights from our household and enterprise census) to compute the study area-wide IRF for each flow component.

For the multiplier, inference is conducted using the wild bootstrap clustered by sublocation, the highest unit of randomization (Cameron, Gelbach, and Miller 2008).[28] We focus on two one-sided hypotheses, namely, that each multiplier is less than zero and each is less than one, as well as tests for the average of the two main multiplier estimates.

4 Tracing out the path of spending

We now turn to tracing out the path of spending induced by the cash transfer experimental intervention. We start by documenting effects for recipient households, then for enterprises and untreated households. Monetary units are USD PPP unless otherwise defined (where the transfer was worth USD 1,871 PPP), flow outcomes are annualized, and monetary outcomes are top-coded at the 99th percentile (as pre-specified), unless otherwise noted.[29]

28. While this procedure may perform poorly in cases where most units in treated clusters are treated and there are few clusters, here at most two thirds of households in the most intensely treated clusters were treated, and there are 84 clusters, far above the 15-20 that MacKinnon and Webb (2018) deem adequate.

29. The main measures were pre-specified, though some groupings vary from the PAP to ease exposition.

4.1 Recipient household effects

The main household expenditure measure is the (annualized) sum of total food consumption in last 7 days, frequent purchases in the last month, and infrequent purchases over the last 12 months.[30] Durables expenditures are the sum of home maintenance, home improvement, and other household durables spending, and the remainder classified as non-durable spending.

As expected, recipient households report significantly higher total expenditure: USD PPP 293 more expenditure than eligible households in control villages (Table 1, column 1), an 11.5% increase over the control village in low saturation area mean of USD PPP 2,537. The estimated total treatment effect, including spatial effects, is larger at USD PPP 338, a 13.3% increase (column 2). This pattern between columns 1 and 2 is a first piece of evidence for localized, positive cross-village spillovers, which is repeated across other outcomes.

The pattern of expenditure effects by category is broadly consistent with earlier work (Haushofer and Shapiro 2016). Both non-durable and durable spending increase substantially. Food expenditure accounts for a sizable portion of the increase in non-durable expenditure in both columns (38% and 59%, respectively). We can reject meaningful increases in reported spending on temptation goods, consistent with Evans and Popova (2017).[31]

Consistent with increased expenditure on durables, asset stocks also increase (Table 1, Panel B). Anecdotally, many recipients withdrew money from M-Pesa immediately and saved via durable assets. The main pre-specified measure of assets includes livestock; transportation (bicycles, motorcycles, and cars); electronics; farm tools; furniture; and other home goods; we add in net household lending to, and borrowing from, both formal and informal sources. This measure of assets increases by USD PPP 182, or 25% of the mean for eligible households in control villages in low saturation sublocations.[32] This measure excludes the values of housing and land, which are harder to measure given thin local markets, but also likely important given existing work shows that households often use GD transfers to spend on housing materials (Haushofer and Shapiro 2016). We separately measure housing value as

30. The survey was quite comprehensive. In addition to food consumption, frequent purchases include airtime and other phone expenses; internet; transport expenses (including petrol); lottery tickets and gambling; clothing and shoes; recreation and entertainment; personal toiletry and grooming; household items, such as cleaning products and candles; firewood, charcoal and kerosene; electricity; and water. Infrequent purchases include house rent/mortgage; home maintenance; home improvements; religious expenses; education expenses; charitable donations; weddings and funerals; medical expenses; household durables, including furniture, lamps, cutlery, pots and pans and other kitchen equipment; and dowry or bride price.

31. While there is likely some under-reporting of temptation goods, the fact that the control group mean is non-trivial demonstrates that at least some households feel comfortable reporting such spending. Given our limited expenditure data immediately after transfer receipt, we cannot rule out that temptation good spending increased temporarily at that time.

32. The mean for eligible households in control villages and low saturation sublocations is USD PPP 724 (with SD 863), less than the overall mean, unsurprisingly since ineligible households are wealthier.

the respondent's self-reported cost to build a home like theirs, and land value as landholdings multiplied by the household's report of the per-acre cost of land of similar quality (in their village). Estimated housing value increases by USD PPP 481, or 79% of the control mean, and estimated land value increases, though this effect is not statistically significant.

Theoretically, the effect of a wealth transfer on earnings is ambiguous: it may reduce labor supply through an income effect, but may also enable productive investment. In the data, recipient households' income from all sources (excluding the GD transfers) does not appear to have decreased: point estimates are positive (USD PPP 78 and 134 in the two main specifications) and the reduced form effect is marginally significant.[33] For labor supply specifically, we do not find that recipient households worked less; if anything, total hours worked by recipient households in agriculture, self-employment and employment increased slightly though not significantly (Table 2, Panel A, columns 1 and 2). This is consistent with the studies reviewed by Banerjee et al. (2017), which generally have found that cash transfers in low and middle income countries do not reduce labor supply.

Interestingly, we observe little heterogeneity in estimated treatment effects (on assets, expenditure, income, and hours worked) among eligible households across eight pre-specified characteristics (Figure A.4), namely, respondent gender, age over 25, marital status, primary school completion, having a child in the household, an indicator for above median measured psychological well-being, and work status (in self-employment or wage employment).

The effect on net transfers received from other households is also notable: the point estimate is negative but not statistically significant, and we can reject large changes in either direction. This suggests that relatively little of the cash transfer was literally shared with neighbors or social contacts.

Overall, these results highlight that cash transfer recipients substantially increased their expenditure on a broad range of goods. This spending was likely financed primarily by the initial transfers themselves, with possibly some contribution from higher earnings. A large share of this spending likely takes place locally: enterprises report that 86 percent of their customers come from within the same village or sublocation. Below we therefore turn to examining impacts on local enterprises.

33. As is common in low-income settings, measured values of consumption are larger than measured household income. Similarly, total measured local area income and firm revenue is lower than expenditures, in part, because measured expenditure includes important categories – including medical and schooling expenses, utilities, rent and mortgage, religious and charitable donations, and dowry, wedding and funeral costs – for which we do not typically measure corresponding revenues in the enterprise data. Expenditure measures may also better capture consumption of own-farm production than the agricultural revenue data.

4.2 Estimating the Marginal Propensity to Consume

The marginal propensity to consume (MPC) sheds light on the inter-temporal decision-making of households, and is an important determinant of the magnitude of a transfer multiplier, as it captures the share of income that is spent—and thus enters the hands of other agents in the economy—rather than being allocated to financial savings or retained in cash (which in our setting might include simply retaining some value on the mobile money platform). The dataset allows us to generate an intuitive estimate of the MPC out of the transfer, obtained conceptually by dividing the total increase in expenditure by the size of the GiveDirectly transfer.[34]

Here we summarize the construction of the MPC in our data; refer to Appendix G for details. An immediate cross-sectional estimate can be obtained by dividing the effect on total household non-durable consumption in Table 1 by the size of the transfer among recipient households. However, this underestimates the MPC, as it is based on consumption as captured in the period preceding household survey administration, with a retrospective timeframe of at most 12 months, compared to the full transfer value, which for many households was distributed at least in part more than 12 months ago. It misses any changes to consumption occurring outside this window, particularly in early months when spending may be the highest. We can improve on this by employing the dynamic regression specification (in equation (7)), which exploits the fact that the timing of survey data collection, relative to transfer disbursement, varied exogenously across households.

Yet this estimate is also a lower bound. As noted above, a limitation of the data collection is the relative lack of household survey data collection in the months after transfers went out, the period when, anecdotally, a large share of the transfer was spent. We thus augment the analysis by making use of data from the closely related Haushofer and Shapiro (2016) study of GiveDirectly transfers provided between 2011–2013 in a nearby part of Siaya County (Rarieda subcounty, lying just outside our study area), which gathered information on household consumption immediately after transfers. The MPC of non-durable goods among recipients in the first three quarters following the transfer there was 0.35 (Appendix Table G.1), consistent with much spending occurring shortly after transfer receipt. Combined with estimates from our data thereafter, recipients' MPC on non-durables over the 27 months post-transfer is 0.64. This implies that most study households receiving cash transfers were hand-to-mouth consumers, allowing us to soundly reject the permanent income

34. This measure is comparable to MPC estimates from tax rebates (e.g., Parker et al. 2013). Alternatively, it may be attractive to divide by the transfer size plus any additional non-transfer income generated over the period. Table 1 shows that this increase is small and not significant, with a point estimate of $\approx 7\%$ of the transfer value. Thus results do not change substantively if this additional recipient income is included, although its inclusion does reduce the estimated MPC somewhat, see Appendix Table G.1

hypothesis in our context.

This estimate still leaves out durable goods expenditure. First note that households report purchasing the vast majority of durables (over 95%) in local shops. These durables may serve as consumption, savings or investment goods. A large share of such purchases in the study sample are consumer durables not primarily intended for productive uses (e.g., radios, furniture). At the same time, formal sector financial savings are limited in rural African settings like ours and much household saving comes in the form of purchases of household durable assets, which necessitates spending on local goods. Thus from the perspective of inter-temporal decision-making, durables are more of a gray area. Yet whether durables are purchased as "savings goods" or "consumption goods", both types of expenditure show up as revenue of local firms and may therefore have similar stimulus effects. Here we rely on the cross-sectional difference in the value of durable assets (including housing) between treatment and control areas among eligible households in our endline data (Table 1). Combining this non-durable expenditure yields our best estimate of the overall MPC in the 27 months following transfers, at 0.93.

Because we are interested in estimating the multiplier effect on local economic activity (within the study area), we next refine the MPC estimate by focusing on spending on local value added, excluding spending on intermediates and final goods produced elsewhere. Spending on goods produced in other parts of Kenya (or the world) does not directly contribute to local GDP (although it could generate multiplier effects at larger geographic scales that we cannot readily assess with our data). We thus derive a bound on the share of spending on local value added. This is closely related to the local degree of openness that features prominently in discussions such as Farhi and Werning (2016). We find that most consumption is in fact of locally produced goods, in line with the well-known fact that a large share of household consumption in rural areas consists of locally produced food and other basic necessities (Deaton 2018). In particular, the enterprise data allows us to bound the share of imported intermediate goods sold in the study area, where, recall, over 95% of household shopping occurs.[35] This conservative methodology yields an upper bound of 18% for the expenditure-weighted share of local non-durable expenditure (and 20% for durables) that may reflect expenditure on imported intermediates, indicating that four fifths of spending is on local value added, and thus that the study area's economy is largely closed.[36]

35. As discussed in Appendix H.1, we determine that at most a fraction $1 - \frac{cost_i + profit_i}{revenue_i}$ of the revenue of firm i is spent on intermediate goods; for each firm type, we then generate a revenue-weighted average upper bound for the share of intermediates in its production function. Next, we make assumptions about what share of intermediate goods is likely imported, conservatively erring on the side of assuming a high share; for instance, we assume that all intermediate goods at clothing stores (which spend up to 38% of revenue on intermediate goods) are imported, which is likely to be an upper bound.

36. In principle this exercise also depends on migration: money spent elsewhere by migrants appears in

Combining estimated import shares with our preferred MPC estimate yields a marginal propensity to consume on local value added, which we denote MPC_{Local}, of approximately 0.76 in this context. Of course, this figure is subject to the data and measurement caveats noted above, as well as assumptions on the share of imports, and so should be seen as speculative. Nonetheless, taking this value of 0.76 to a basic static Keynesian model, the transfer multiplier effect on local output would be $\frac{MPC_{Local}}{1-MPC_{Local}} \approx 3.2$. We dynamically estimate the multiplier using all household, enterprise and price data in section 5 below.

4.3 Enterprise effects

There are large increases in revenue for enterprises in both treatment and control villages (Table 3, Panel A). Revenues in treated villages increased by USD PPP 357 per household, a 47% increase, while those in control villages increased by USD PPP 244 (32%). Revenue gains are concentrated in the retail and manufacturing sectors: both treatment and control villages experience statistically significant increases in manufacturing revenue of similar magnitudes – USD PPP 109 and 110, respectively – while treatment villages see larger gains in retail revenue (USD PPP 141 versus USD PPP 65, Appendix Table A.2).

Estimated effects on profits are positive, but moderate in magnitude and not significantly different from zero. In fact, profit margins (measured as the ratio of profit to revenues) fell (Table 3, Panel A, Row 5). We also see no evidence of firm entry, as one might have expected if enterprises were becoming more profitable (Panel C). Overall, the data indicate that higher revenues were largely absorbed by increased payments to various factors of production. While we do not observe all of these payments, we do see significant increases in the factors that we directly measure, and particularly the wage bill: enterprises in treated (control) villages increase spending on labor by USD PPP 82 (71), a sizable change relative to the mean.

Strikingly, we do *not* see strong evidence of a firm investment response. Estimated increases in fixed capital investment are small, and we can reject large changes (Panel B, Row 2). We do see a modest increase of USD PPP 35 in inventories for enterprises located in treated villages, yet even this appears to be less than proportional to the increase in firm sales; in other words, these enterprises are, if anything, operating leaner business models (Panel B, Row 1). This pattern of results suggests that the expansion in enterprise activity is driven more by the shock to local aggregate demand than by a relaxation of credit constraints that had previously limited investment.

---

our data (as we tracked and surveyed them) but does not contribute towards the share spent on local value added. In practice household migration was uncommon, with 5% of control low-saturation household migrating, and unaffected by treatment (Table A.9, Row 1). Estimated treatment effects among non-migrants are also essentially identical to overall average effects (Table A.9, Panel 2).

One caveat to this point is that some household assets are difficult to categorize into "productive" assets as opposed to consumer durables. For example, bicycles may be used for personal transportation (i.e., to visit friends), but could also be used as a bicycle taxi to generate income. We therefore inclusively categorize as "potentially productive" both livestock as well as a number of non-agricultural assets that could potentially be used for income-generating activities (beyond simply renting out the asset).[37] When we do so, overall roughly half of the increase in household asset ownership documented above is in what we believe to be purely non-productive assets, with small gains in productive agricultural assets (e.g., farm tools) and a modest gain for potentially productive assets (Table A.1). We also fail to detect any investment response for non-agricultural enterprises owned by recipient households: neither investment nor inventories increase relative to eligible owners in control villages (Table A.3, Panel B). Taken together, these patterns are also consistent with the cash transfer program generating only a limited local investment response.

4.4 Non-recipient household effects

There are positive and significant expenditure effects for non-recipient households. Column 3 of Table 1, Panel A presents results based on Equation (3). Notably, the magnitude of these gains (USD PPP 334, p-value < 0.01) are quite similar to those of recipient households (USD PPP 338). The pattern of expenditure increases is also broadly similar to that for recipient households, except that spending on durables does not increase among non-recipient households. One possible reason for the similarity in overall spending impacts is that the timing of effects on recipient and non-recipient households may be different, with recipient households showing impacts earlier than non-recipient households, but effects converging by roughly one year after the final transfer was received. A further potential mechanism is that labor earnings increase differentially: among non-recipients, annual labor income increases by USD 229, while the figure is USD 134 for recipients. For wage earnings, the figures are USD 183 and USD 73, respectively. Thus, the similar impacts on expenditure among recipient and non-recipient households may partly be explained by a lower labor income response among the former. Finally, note that non-recipient households include both eligibles and ineligibles, and, as shown in Table A.8, most of the gains accrue to ineligibles. These comparatively wealthier households might be gaining more from business and additional labor income, and may be imperfectly substitutable with eligibles in the labor market. As a result, they may experience a larger increase in wages than recipient and non-recipient eligibles.

37. Potentially productive non-agricultural assets include bicycles, motorcycles, cars, boats, kerosene stoves, sewing machines, electric irons, computers, mobile phones, car batteries, solar panels or systems, and generators. Examples of residual non-productive assets include radio/CD players, kerosene lanterns, beds, mattresses, bednets, tables, sofas, chairs, cupboards, clocks, televisions, and iron sheets.

How did non-recipients fund these consumption gains? One possibility is that they are dis-saving, perhaps due to social pressure to "keep up with the Joneses", their neighbors who received the transfer. However, this does not appear to be the case: estimated treatment estimates for total assets, housing and land values are all positive, although not all are significant (Table 1, Panel B). Nor do we observe a borrowing response for non-recipient households from either formal and informal sources (Table 2, Panel C, column 3). A second potential explanation is that expenditure gains reflect inter-household transfers to non-recipient households, as documented in Angelucci and De Giorgi (2009) for Mexico. This also does not seem to be the case, as we find no significant increase in net transfers received by non-recipient households, and the point estimate of USD PPP 8.75 is less than 3 percent of the expenditure gain for non-recipient households; this mirrors the lack of an effect on net transfers among recipient households noted earlier.

Rather, the data suggest that consumption gains are driven by higher earned income: total annualized income increases by USD PPP 229. It is often argued in development economics that survey estimates of consumption are better measured and often substantially larger than estimates of income, particularly for poor households (Deaton 2018). While this is true in our case, we cannot reject that the total effect on income is the same as the effect on consumption expenditure for non-recipient households ($p = 0.23$). Income gains come largely from wage earnings, which increase by USD PPP 183, with a smaller and not significant contribution from profits from owned enterprises. These results are broadly in line with the enterprise results, in which profit increases were modest and marginally significant while the wage bill expanded significantly, by 68 and 58% in treatment and control villages, respectively (Table 3, row 4). Higher wage earnings appear more likely to reflect higher wages than increased labor supply, as the point estimate for overall household labor supply is actually somewhat negative (although there does appear to be an increase in respondent hours worked for wages, Table A.4). Hourly wages earned by non-recipient household increase meaningfully, although the estimate is only marginally significant (Table 2, Panel A).

To sum up the results so far, cash transfer recipient households receive and spend most of the transfer, leading to higher local enterprise revenues. This positive aggregate demand shock, in turn, appears to increase the income of local non-recipient households, leading to higher spending on their part. This pattern provides initial evidence for a positive multiplier effect of the cash transfer program, an issue we return to below.

4.5 Effects on output prices

We turn next to effects on consumer goods prices in order to understand the extent to which other monetary impacts are real as opposed to nominal. Overall, we find small, positive and precisely estimated effects on consumer goods prices. For our overall expenditure-weighed log-index of market prices both the ATE and average maximum transfer effect are small and precisely estimated near zero (Table 4). The tight standard errors allow us to rule out even relatively small price effects: with 95 percent confidence, the ATE across the study period is below 0.0022 log points, or 0.22 percent. For the average maximum transfer effect across markets, the upper bound of the 95 percent confidence interval is 0.01 log points, or 1 percent. Price effects are also small across almost all product categories. In particular, food prices are in line with the overall price index, and durable prices do not increase meaningfully. To help mitigate concerns that results may be sensitive to the price index weights or product classification, we find that average price inflation is below 1.2% for every product (Figure A.6; for alternative specifications and product classifications, see Appendix E.2).

Variation in price responses is generally in line with theoretical predictions. We observe somewhat larger price increases in markets less integrated into the local economy. Columns 3 and 4 split markets into those above and below median market access, with estimated effects typically more positive in more remote markets. Figure A.5 further breaks this pattern down by quartile of market access, with lower values reflecting more isolated markets. Panels A-C show a small amount of inflation for less tradable goods only in the most isolated markets, and smaller and less precisely estimated effects for more tradable goods, with less of a clear pattern across market access quartiles. Inflation for less tradable goods in isolated markets nonetheless remains limited, at 0.2-0.3% on average. We also carried out enterprise phone surveys of a subset of enterprise types during the period in which transfers were going out, which collected price data on a limited number of products; inflation for these local manufacturing and services prices is also limited (see Appendix E.3).

These patterns are qualitatively similar to findings in Cunha, De Giorgi, and Jayachandran (2018), who study the price effects of an in-kind food and cash transfer program in Mexico (where the household income shock was similar in magnitude to the Kenya program we study): in-kind transfers there lead to price decreases, while cash transfers lead to price increases, but their estimated effects are small except in remote villages. Filmer et al. (2018) estimate inflation of 5 to 7% for protein-rich foods in the Philippines, with smaller effect for other product categories. Burke, Bergquist, and Miguel (2019) show that a credit intervention impacting the supply of staples also affects local grain market prices in a different Kenyan region. Reconciling these results with ours is a task for future research.

4.6 Effects on input prices

We next examine effects on the prices of major factors of production: labor, land and capital. Table 2 presents estimated effects on these prices measured in the household survey data.

We find some evidence of higher wages. In row 1 of Table 2, we examine wages for employees using household survey data.[38] In the reduced form specification, eligible households in treatment villages earn USD PPP 0.11 more per hour, on a base of USD PPP 0.70. This effect is no longer significant, however, when we also estimate across-village spillovers. For non-recipient households, the increase is even more marked at 0.19 USD PPP per hour, and significant at the 10% level. These potentially large wage effects do not seem to be driven by large labor supply responses. In row 2, we calculate the total hours worked by adult household members in agriculture, self-employment and employment, and estimate effects at the household level. Effects are relatively small and not significant. Together with the fact that enterprise wage bills increased, these patterns are strongly suggestive of positive local wage effects (Table 3). This in turn suggests that labor markets in this area are fairly localized, at least over the time horizon we study, which is consistent with the fact that we see little evidence of impacts on measures of migration (Table A.9). In the longer run, labor may become more mobile, helping to equilibrate any induced wage differentials.

Effects on estimated land prices are positive and economically meaningful (at 9-14%), but not significant (Table 2, Panel B). Since our measure of land prices is a somewhat noisy one—formal sales are rare so we use respondents' self-reports of the amount per acre land like theirs in the same village would sell for—we also examine land rental prices as a robustness check, which yields data on actual land transactions for a subset of respondents. We do not find significant effects on land rental prices (Table A.5). Unsurprisingly, given land should be in relatively fixed supply in the short-run, we find little change in total landholdings among recipient households or those in more heavily treated areas. We also find no effects on total land rentals, nor on the total amount of land used for agriculture (Table A.5).

We estimate fairly precise null effects on interest rates and total borrowing (Table 2, Panel C), where we measure household borrowing from both formal (e.g., banks, mobile credit services) and informal (moneylenders, family and friends) sources. The loan amount reports total borrowing across sources in the last 12 months, setting those who did not borrow equal to zero. Note that the loan-weighted interest rate is the monthly interest rate on the most recent loan by source, weighted by the total amount of borrowing (by source); we include informal loans without interest, which brings down the average rate.

5 The transfer multiplier

We next examine what the household and enterprise responses imply for the aggregate level of economic activity, and specifically for the value of the local multiplier of cash transfers,

38. We include all household members that report working for wages, and calculate their hourly wage based on hours worked in the last 7 days and their monthly salary (adjusted to weekly scale).

where 'local' refers to the entire study area. We define this multiplier $\mathbb{M}$ as the cumulative effect of transfers on local real GDP, relative to the total amount T transferred in real terms, over a given time interval:

$$\mathbb{M} = \frac{1}{T}\left(\int_{t=0}^{t=\bar{t}} \Delta GDP_t\right) \tag{8}$$

The size of the transfer multiplier is generally thought to depend in part on the policy context in which outlays are made, and in particular on the extent to which (i) monetary policy reacts, and/or (ii) households and firms expect levels of current or future taxation to change. Our setting is unusual in a useful way: because we observe a large one-time fiscal outlay that was made philanthropically, funded from outside of the economy we study, and small relative to the overall Kenyan economy, we can reasonably expect to measure a "pure" external transfer multiplier that should be independent of such effects. This feature generates estimates that can be thought of as a model primitive, and with which estimates from other financing scenarios can be contrasted.

As noted in Section 4.2, an initial calculation of the transfer multiplier as $\frac{MPC_{Local}}{1-MPC_{Local}}$ suggests that it may be substantial, at around 3.2. In this section, we refine this estimate by both accounting as fully as possible for effects on all components of GDP, including spillover effects, and accounting for dynamics. To get at real values, we deflate all monetary outcomes and transfer values to January 2015, linking the overall monthly market price index in the nearest market to each observation (Appendix H presents a nominal version).

Following national accounts definitions, the expenditure-based measure of local GDP is $GDP_t = C_t + I_t + G_t + NX_t$, where C_t is consumption expenditure on non-durables and durables, measured as quarterly consumption plus accumulated assets and housing stock at endline.[39] To avoid potential double-counting, we exclude expenditure on home durables, home improvements and maintenance from the consumption expenditure measure as part of this expenditure may be reflected in an accumulation of assets. In addition, we exclude net lending as well as land values from the asset measure because changes in land values may not be driven purely by investment, and because we think of land supply as being essentially fixed. We exclude local government expenditure, G_t, as Walker (2018) shows that the intervention had a precisely estimated null effect on it.

Since we also measure household and enterprise income, we can construct a dual income-based measure of local GDP as the sum of factor payments and profits: $GDP_t = W_t + R_t + \Pi_t + Tax_t - NFI_t$, where W_t is the total household wage bill, R_t are rental expenses of

39. Note that by measuring impacts on asset stocks we (correctly) do not count transfers of existing assets between local agents as GDP, since these increase one agent's balance sheet while decreasing another's. Such transactions only introduce bias if they involve a non-local counterparty, as discussed below.

local enterprises (assuming those are paid to capital owners within our study area), Π_t are enterprise profits, and Tax_t is total enterprise taxes.[40]

For flow variables, we follow Section 3.8 to generate IRFs, which we integrate over time. For two components of GDP, we are instead able to measure impacts on the integral of flows over time by simply measuring impacts on accumulated stocks, simplifying the problem. Specifically, we measure effects on durable consumption expenditure using effects on the stock of endline household durable goods and the value of respondents' home, and effects on inventory investment using effects on current inventory stocks at endline. One drawback is that these figures are likely to under-estimate cumulative spending to the extent that some assets depreciated between the time of purchase and measurement, although over the limited timeframe considered this may be a second-order concern; any such bias would tend to reduce the estimated multiplier. In the graphical presentation, we assume that any effects on these stocks occurred equally across all post-treatment quarters.

Overall, we view the expenditure- and income-based multipliers as two distinct measures of the same underlying concept, each with its own limitations. Reflecting this, below we estimate them jointly and test individual as well as joint hypotheses across the two measures. We discuss limitations and robustness in detail in Section 5.2, including adjustments to account for the fact that we do not directly observe NX_t or NFI_t.

5.1 Multiplier estimates

We estimate a sizeable multiplier using both main approaches, in line with the back-of-envelope figure derived above. The estimated expenditure multiplier is 2.53 (Table 5, Panel A). 46% of this effect is driven by consumption expenditures. Household asset purchases and enterprise investment make up another 32% and 19% respectively, and enterprise inventories are not quantitatively important. While part of the asset response could potentially reflect productive investments by household-operated enterprises, at least 43% of the asset response comes from non-productive assets, across both recipient and non-recipient households (see Table A.1). Taking this into account, consumption alone leads to an estimated multiplier of at least 1.5, underscoring the overall point that cash transfers appear to have led to a predominantly demand-side driven increase in local economic activity.

The estimated income-based multiplier is quite similar in magnitude to the expenditure-based multiplier, at 2.28 (Panel B), and we cannot reject that they are the same ($p =$

40. We employ the household rather than enterprise wage bill, as the household survey sample is larger and includes individual-level wage earnings data. We omit land rental income because we do not see any significant evidence of effects on this above. In principle, a third approach to estimating GDP would be to aggregate value added from local enterprises; we do not implement this as we did not collect sufficiently comprehensive data on enterprise expenditures on intermediate inputs.

0.88). This is notable since it is calculated using a completely distinct set of component measures. Of this total effect, we find that 64% reflects increased enterprise profits, 30% increased wages, and a much smaller contribution comes from from capital income and taxes taken together. As noted above, the increase in consumption, and the smaller increase in investment, is therefore primarily accounted for by higher profits and wages. Of course, in our context of predominantly single-person firms, "profits" likely reflect some mix of true economic profit along with returns to the owner's capital and labor inputs. Regardless of the exact mix, however, this sum should be appropriate for our goal of calculating the aggregate income-based multiplier.

When we examine the relative contributions of recipient and non-recipient households to both multipliers (as shares of the total household contribution), we find that non-recipient households account for 80% of the household contribution to the expenditure multiplier and 85% of the contribution to the income multiplier, both of which are somewhat higher than their share in the local population of 67%. This suggests that analysis focusing only on recipient households may be missing sizable shares of program effects.

An advantage of this "macro-experimental" approach to estimating the multiplier is the ability to conduct statistical inference. To start, we reject the null of a negative multiplier (with a value less than zero) at the 10% level using either approach (Figure 1 and Table 5), and reject the null at $p = 0.02$ when testing the joint restriction, and at $p = 0.04$ when testing the average of both multipliers. Since the two measures exploit distinct data, we gain statistical power by examining both measures together. Rejecting a negative multiplier on real GDP is important, ruling out, for example, that prices adjusted immediately to increased spending, netting out any real effects. Testing the null hypothesis of a multiplier less than one has been a central goal of recent research on the fiscal multiplier, since it would imply a crowd out of private spending, but this value does not have the same interpretation for transfer multiplier estimates like ours where all spending is private. Nonetheless, rejecting a transfer multiplier value less than one constitutes a conservative test for the existence of positive output spillovers for non-recipients, holding even in the extreme case of $MPC = 1$.[41] Using the expenditure or income-based approach alone, the p-value on this test is 0.14 and 0.23 respectively. The average of the two multipliers is 2.40 ($SE = 1.38$, $p = 0.15$), and the hypothesis that the multipliers are jointly less than one is rejected at the marginally significant $p = 0.07$ level.

Figure 1 presents these results graphically, breaking up the aggregate multipliers into

41. A less conservative test of the no-spillovers hypothesis would be to test if the multiplier estimate is less than the MPC of recipients (which is strictly less than 1). Because this MPC is itself somewhat imprecisely estimated in our data (see section 4.2), this approach does not necessarily increase power.

quarters after transfers went out. Panel A presents the expenditure-based multiplier. The increase in GDP is fairly stable over time; in fact, we cannot reject that the expenditure response is constant across all quarters (p-value of 0.73). It increases slightly up to a peak after 9 months (when the second lump-sum transfer has been received), and then slowly declines. Interestingly, we reject a null effect as late as two years after the transfer, suggesting that the true multiplier (out to an infinite time horizon) could be larger still and that our estimates are likely to be lower bounds. The less precisely estimated effects (with larger confidence intervals) during the first three quarters afters transfers go out are visually apparent. The income multiplier, on the other hand, visually appears to fluctuate more over time (Panel B): it is marked by a strong early response in profits, while wages appear to take longer to rise. Yet as with the expenditure measure, we cannot reject equality of all quarterly coefficients (p-value of 0.76).

These estimates are somewhat larger than the higher end of recent fiscal multiplier estimates in the context of public spending in the United States (Chodorow-Reich 2019; Nakamura and Steinsson 2014), where they tend to range from 1.5 to 2.0. As noted above, the magnitudes of transfer versus fiscal spending multipliers are not directly comparable. The differences between our results and existing estimates may also reflect the relative levels of economic development and other structural differences between the Kenyan and US economies (such as the degree of openness of local economies, the share of hand-to-mouth consumers and the existence of financial savings opportunities), differences in data and measurement, as well as any effects on (or expectations of effects on) either monetary policy or future taxes in the US, the latter being response effects that this study's experimental design usefully allows us to avoid.

5.2 Alternative assumptions

The expenditure- and income-based measures of GDP we generate are based on unusually rich underlying data, but each has potential limitations. In particular, each may misattribute transactions between agents located in the study area and counterparties located outside.

In the expenditure case, the main concern is that we do not directly observe net exports (NX_t). Imports show up as expenditure but are not local GDP, while exports do not show up in expenditure but are part of local GDP. To the extent that cash transfers decrease (increase) net exports from the study area, our expenditure multiplier would overstate (understate) the multiplier. Intuitively, we might expect net exports to fall following a large external income transfer: since many local firms are retail establishments, imports of intermediate goods (including packaged consumer goods ready for sale) would likely increase. This suggests that the expenditure-based approach might be upwardly biased.[42] Note that transactions between agents *within* our study area are correctly accounted for: for example, if study village A imports goods from study village B then the value of these goods should be included in local GDP as they are produced within our study area. Of course, increases in net imports could in part reflect increases in economic activity outside of the study area due to the cash transfers, which our concept of the local multiplier does not capture but which are a part of the broader impact of the intervention.

As a robustness check to gauge the magnitude of the potential bias in the expenditure-based measure due to imports of intermediate inputs, we first assign each component in the non-durable and durable expenditure measures to enterprise types at which the good is most likely to be purchased (using revenue shares of different enterprise types, where appropriate). As noted in Section 4.2, this conservative methodology yields an upper bound of 20% of local spending that may reflect expenditure on imported intermediate goods. If imports scale linearly with expenditure, this suggests a transfer multiplier of at least 2.01 on local expenditure alone (see Appendix H).

In the income case, potential bias could arise if there are changes in net wage income (NFI_t) earned outside the study area, since this is not considered part of local GDP. This bias seems unlikely to be quantitatively important in our setting: 86% of all non-farm employees are family labor (and therefore presumably overwhelmingly local), and among individuals employed for a wage, only 6% report an employment contact address outside the study area. To the extent some bias remains, we would expect it to be negative (towards zero), if net labor income earned outside the study area decreases in response to higher local business revenue, employment and wages. This suggests that the income-based approach may yield a lower bound on the multiplier. Consistent with this bounding logic, the estimated income multiplier is somewhat smaller in magnitude than the expenditure multiplier.

We also consider several alternative multiplier estimates that treat the first three quarters of data post-transfer, as well as prices, in different ways. A conservative approach excludes effects on GDP during the first three quarters after transfers arrive, which may be statistically attractive in a mean squared error sense since it yields more precise, if surely somewhat downward biased, estimates. Under this assumption, the expenditure (income) multiplier estimate is 2.04 (1.45), which is smaller than the preferred estimates in Table 5, as expected, and estimates attain greater statistical significance (see Appendix Table H.4). An arguably

42. Note that direct imports by households themselves are unlikely to increase because on average only 10% of households report ever shopping at a market outside our study area, and overall the impacts we see on household spending and local enterprise revenue are fairly similar, suggesting that consumer spending was quite localized. Similarly, non-farm businesses report only 5% of customers coming from outside the study area, and that share does not change significantly in response to treatment.

more realistic method utilizes the household consumption data from Haushofer and Shapiro (in prep.) for the first three quarters post-transfer, as we did in the construction of the MPC discussed above. This yields a larger estimated expenditure multiplier of 3.09, with the increase due to a greater contribution from recipients' non-durable consumption (see Appendix Table H.4). A final alternative presents the multiplier in nominal rather than real terms: the nominal expenditure (income) multiplier is 2.66 (2.55), see Appendix Table H.5. Given our quantitatively small price effects, the differences between these and the real estimates presented in Table 5 are mainly due to the moderate degree of overall price inflation during the study period.

6 Welfare implications

Transfer multiplier estimates have typically been used for positive economic analysis, to predict how fiscal policy will affect output. Yet since output is not social welfare, how fiscal policy affects welfare is a distinct issue. Classic derivations of fiscal multipliers from accounting relationships such as the "Keynesian cross" could not deliver parallel statements about welfare as they were not grounded in models of individual preferences. While recent papers have focused primarily on estimation (Ramey 2019), a few have examined the relationship between fiscal multipliers and welfare in the context of micro-founded models, emphasizing that multipliers need not be sufficient statistics for welfare or (consequently) for optimal policy (Sims and Wolff 2018). In fact, Mankiw and Weinzierl (2011) construct examples in which the interventions with the largest multipliers have the *least* impact on social welfare. The program evaluation literature on cash transfers, meanwhile, has largely focused on estimating behavioral responses without exploring what these mean for welfare.

Here we examine the broad channels through which transfers could affect household welfare, and how these relate to the transfer multiplier. Let indirect utility function $v_i(T_i, T)$ define the utility achieved by household i when it receives a (possibly zero) transfer T_i while other eligible households in the area receive transfers of T each. We are interested in characterizing how T affects the quantity T_i^* defined by $v_i(T_i^*, 0) = v_i(T_i, T)$, in other words, the transfer that would make household i indifferent between receiving T_i^* on the one hand, and experiencing the intervention we study on the other. Notice that if there were no general equilibrium effects, in the sense that v_i did not depend on T, then we would simply have $T_i^* = T_i$, i.e., the tautology that the value of receiving a dollar is a dollar.

We think of v_i as the value of some generic underlying optimization problem

$$v_i(T_i, T) = \max_{x_i} u_i(x_i, x_{-i}(T)) \text{ s.t. } x_i \in X(T_i, T) \tag{9}$$

Here u_i represents preferences over variables x_i which the household chooses from a set X, as well as variables x_{-i} chosen by others. This formulation delineates two ways in which T can affect the utility of household i. First, it may change market outcomes that determine the choice set X – for example, prices or income from various sources. We therefore need to interpret the impacts on output that generate the transfer multiplier through this lens. Second, it may change behaviors $x_{-i}(T)$ that affect i's well-being without appearing in the transfer multiplier (e.g., through externalities).

6.1 Market outcomes

An increase in (real) output must reflect some combination of (i) an increase in the *employment* of factors of production and (ii) an increase in their aggregate *productivity*. While the latter represents an unambiguous welfare gain, the former comes at an opportunity cost – the value of foregone leisure, for example, in the case of labor inputs, or of foregone present consumption in the case of capital inputs. Appendix J provides a formal illustration of the mapping from household welfare to aggregate output, emphasizing this point. The discussion also illustrates how household welfare differs from household expenditure, which is often used in the program evaluation literature as a proxy for well-being. Specifically, expenditure does *not* take into account the opportunity cost of supplying labor (or other inputs), and over any finite time interval incorrectly interprets dis-saving as a welfare gain.

A key question is thus the extent to which the output response we observe can be explained by increases in the supply of scarce factors of production. In the data, we find fairly limited evidence of increases in the employment of either land, labor, or capital. Land is in relatively fixed supply; agricultural households do not report owning or renting more of it (Table A.5) and we would not expect it to be a limiting factor in the sectors in which the output expansion is concentrated (namely, retail and manufacturing). Total household labor supply does not change significantly (Table 2), though we do see a net shift out of self-employment and into wage employment (Table A.4, Panel A), with the latter increasing by 1.8 hours per person per week on average across recipients and non-recipients. These estimates are not statistically different from zero, however, and even under generous assumptions can explain only around a 5% increase in real output, well below the observed response.[43]

As for capital, the non-agricultural enterprises that increased their output did not increase investment in fixed capital (Table 3, row 7) and, while increasing inventories somewhat, actu-

43. Specifically, an increase of 1.8 hours per person is a 7.7% increase in wage labor hours. Assuming a Cobb-Douglas production function with a labor share of 2/3, and no productive value of time given up from self-employment, this implies a 5.2% increase in real output.

ally decreased them slightly in proportion to sales (Table 3, row 6). Moreover, if investment were driving output increases then we would expect to see these increases concentrated in enterprises owned by recipients, who gained access to a new source of capital, but if anything we find the opposite (Table A.3). Overall, the limited factor supply response suggests that the bulk of the output response we estimate must be attributable to productivity gains, and should thus be valued at roughly \$1 per \$1 in welfare terms. (We discuss productivity further in Section 7.)

The distribution of benefits also matters for welfare to the extent we value more highly expansions in the budget sets of poorer households. While transfers were targeted to relatively poor households, we have seen that large spillovers accrued to their somewhat richer neighbors. Indeed, we find no significant reductions in village-level Gini coefficients for consumption expenditure or wealth in treatment villages, and a small and marginally significant ($p < 0.1$) *increase* for wealth in control villages (Table A.7). We also reject in most cases the null that observed effects on Gini coefficients are equal to the counterfactual changes we might have expected had there been *no* spillovers. Overall, the patterns underscore the large spillover gains for non-recipient households: wealthier non-recipients benefit along with recipients, on some dimensions so much that inequality may slightly increase.

Distributional effects could also work through prices; while the overall price level changed only slightly, changes in relative prices could transfer value between net buyers and net sellers of goods and services. However, as for the overall index, effects are muted across all individual goods prices that we measured, with nearly all changes within a -1% to +1% range, indicating that any redistributive effects via price changes are likely to be very small (Figure A.6).

6.2 Non-market outcomes and externalities

We measured several outcomes that do not enter into our multiplier calculation but that arguably influence well-being or proxy for it, and may thus capture externalities either between or within households ($x_{-i}(T)$ in Equation 9). Specifically, we examine indices for psychological well-being, health status, food security, education, female empowerment, and security from crime. Each index is the inverse-covariance-weighted sum of component z-scores signed so that positive values indicate better outcomes.[44] The index for psychological well-being can be interpreted as a measure of overall well-being. The next four indices arguably capture intra-household externalities, while security from crime is an inter-household externality.

44. The first five of these were pre-specified as primary outcomes; the components of the security index were pre-specified as part of a family of outcomes, though combining them into an index was not. Details on index construction and results for components are in Appendix A, and PAP details in Appendix I.

For recipient households, we find positive and significant reduced-form effects for four of the six indices: psychological well-being, food security, education and security. Estimates are close to zero and not significant for health and female empowerment.[45] Total effects including spillovers are similar for all but the security index. For non-recipient households, on the other hand, we find no significant effects except for a 0.1 SD increase in the education index ($p < 0.10$). We do not find evidence of adverse spillover effects for non-recipient households on any index, with point estimates positive for all but the security index, which is indistinguishable from zero (-0.02 SD, SE 0.07). Village public good provision was also unchanged (Walker 2018).

Overall, this pattern of findings suggests that the most important welfare effects were market-mediated, though of course there may be other external effects we did not measure. A possible exception is the impacts on inequality noted above: to the extent that households care about comparisons with neighbors, these may constitute a form of "psychic externality."

7 Discussion: utilization of productive capacity

The results raise the question of which features of the local economy enabled it to respond elastically to a large aggregate demand shock. While fully addressing this is beyond the scope of the present project, we outline what can be said given available data.

Any explanation of these patterns must apply to the retail and manufacturing sectors specifically, as it is here rather than in agriculture or services that output gains are concentrated (Table A.2). Moreover, it cannot rely on an increase in the *employment* of factors of production, since we find little evidence of this (Section 6.1). It must instead reflect an increase in the *utilization* of factors employed, as well as in the throughput of intermediate goods. This notion is consistent with our observation (during fieldwork) of the retail and manufacturing enterprises in the area, which typically involve some degree of "on-demand" production. A retail establishment, for example, requires premises and an employee to "mind the shop," but once these are in place the volume of goods it sells depends largely on consumer demand. Similarly, many small-scale manufacturing enterprises require equipment and staff to be in place but then produce only when customers arrive. In fact, about 60% of manufacturing revenue accrues to just two enterprise types, grain mills and welding shops, both of which largely operate in this way.

These examples suggest retail and manufacturing sectors in which there are important inputs whose costs are fixed over the relevant ranges – e.g., a building, milling machinery, or hiring an employee – and whose utilization thus depends on demand. While we did

45. The latter (non)-result contrasts with Haushofer and Shapiro (2016) who found increases in female empowerment and reductions in domestic violence among households receiving a similar transfer.

not measure capacity utilization directly, some indirect evidence suggests the existence of meaningful slack. The average non-agricultural enterprise saw just 1.9 customers per hour, in between which other inputs (i.e., employee time, fixed capital, inventories, etc.) may sit idle. For labor inputs in particular, 72% of non-agricultural enterprises have just a single employee, which suggests that (due to integer constraints) the labor input is essentially fixed over the relevant range.[46]

Given this structure of production, we would expect the revenue from additional sales to be paid out to the suppliers of intermediate goods, the suppliers of elastic factors of production (whose marginal product increased as they became better utilized), and to enterprise owners to the extent they are able to extract economic profits. We do not directly measure purchases of intermediates, but upper bounds for the expenditure-weighted share of intermediate inputs in total sales are sizeable, at 58% in the retail and 18% in the manufacturing sector (see Appendix H). We also see an increase in wage bills, which accounts for 26% of increased revenue (Table 3). Estimated effects on profits, meanwhile, are positive but modest and not statistically significant (and may in any case be better interpreted as returns to the owners' capital or labor which, as usual, are difficult to distinguish from true economic profits).

While suggestive, this interpretation of the supply side response to a demand shock is consistent with other recent findings. In Uganda, Bassi et al. (2019) find that employees in on-demand manufacturing (e.g., welding, furniture-making) spend about 25% of time "waiting for customers" or "eating and resting." More broadly, it relates to the old idea in development economics that it might be possible to expand production without notable price inflation due to the availability of slack capacity. Classic arguments focused on "surplus labor" due to artificially high wages (Lewis 1954), while here both labor and capital appear to have been underutilized due to limited flexibility to scale their employment to match demand. Local mechanisms to address this through better coordination, such as periodic markets, do so imperfectly, leaving some degree of residual excess capacity.[47]

A deeper question is whether specific market failures contribute to slack capacity in steady-state in rural Kenya. Here we speculate about possibilities and directions worthy of future investigation. The most immediate explanation revolves around the small scale of local market activity; for instance, a single grain mill typically serves each village. While the capacity provided by the standard grinding machine and the worker staffing it may generate

46. Note that while we do not observe a large response in reported labor supply, we are not measuring utilization of labor capacity at the intensive margin; e.g., we do not distinguish between the time a shopkeeper waits for customers or serves them.

47. A growing literature also finds evidence of excess capacity in rich countries, especially in periods of recession (Murphy 2017; Michaillat and Saez 2015; Chodorow-Reich 2019).

positive profit for the owner, this capacity may also exceed average local demand, implying excess steady-state capacity which could be engaged following a demand shock. The small scale of local markets is itself likely to reflect the poor road quality and high transport costs that characterize rural Africa (Foster and Briceno-Garmendia 2010). The same logic suggests that multipliers could be smaller in cities due to their greater population density and better transportation infrastructure, not to mention the fact that rural economies are often relatively closed, with large shares of consumption coming in the form of locally produced food and other basic necessities.

Contracting frictions and institutions may also affect local market structure and capacity. For instance, Bassi et al. (2019) document a pattern of small industrial clusters in neighboring Uganda, in which a dozen small carpentry firms producing nearly identical products may co-exist in the same area. Each of these separately owned firms has one or at most a few employees, and they are characterized by the slack labor capacity noted above. Consolidation into fewer, larger firms – each better utilizing workers' time and any installed machinery, and run by a more capable manager – could conceivably reduce slack and free up labor to shift to alternate activities. Further research on the legal, financial and output market frictions that prevent horizontal integration of this kind would be useful.[48]

8 Conclusion

A large-scale cash transfer program in rural Kenya led to sharp increases in the consumption expenditures of recipient households, and extensive broader effects on the local economy, including large revenue gains for local firms (that line up in magnitude with household consumption gains), as well as similar increases in consumption expenditures for non-recipient and recipient households approximately a year and a half after the initial transfers. Firms do not meaningfully increase investment, and there was minimal local price inflation, with precisely estimated effects of less than 1% on average across a wide range of goods. Two independent calculations of the local transfer multiplier using consumption data and income data yield estimates of approximately 2.4, and reject the hypothesis that the multipliers are less than or equal to 1 with 90% confidence. Several suggestive patterns are consistent with the existence of "slack" on the production side in our context, which may partially account for the large estimated multiplier.

Concerns that cash transfer programs like the one we study could have adverse consequences for non-recipients were not borne out in our setting. Firm revenues and non-recipient households' consumption expenditures rise substantially in areas receiving large cash trans-

48. Another fruitful direction for further investigation is whether cash transfers triggered a productivity-enhancing re-allocation of factors of production across sectors and firms (Hsieh and Klenow 2009).

fers; there is little price inflation; overall economic inequality does not increase meaningfully in treated areas; nor are there negative effects in terms of domestic violence, health, education, psychological well-being, and local public goods. Instead, the positive spillovers we find suggest that RCTs of cash transfer programs that simply compare outcomes in treatment versus control villages may understate true overall impacts by ignoring the general equilibrium effects that we capture (along the lines that Miguel and Kremer (2004) argue in the context of a health program).

This study is among the first to exploit randomized controlled trial methods to directly estimate macroeconomic parameters and more broadly capture large-scale aggregate effects of a development program. The multiplier effects that we focus on here have been the subject of intense interest since at least the seminal work of Keynes (1936). Our approach thus provides a novel counter-example to the well-known critique that RCT methods are not well-suited to studying the "big" questions in development economics (Bardhan 2005; Easterly 2006; Deaton 2010). We demonstrate that there need not always be a trade-off between a study's rigor and its relevance: economics research can increasingly achieve both (Muralidharan and Niehaus 2017; Burke, Bergquist, and Miguel 2019).

The extent to which the multiplier results apply to other settings merits further discussion. They are likely particularly relevant for low or middle income economies that share structural and institutional features with Kenya, including many other African settings. One open question is the extent to which targeting of particular types of households, and the distribution of spending propensities across households, affect the multiplier: for example, spillover effects might have been more muted if the program had also targeted transfers to some better-off households with lower marginal propensities to spend on local goods than the poor rural households we study. A second issue is how the multiplier may vary over the business cycle. It is noteworthy that we estimate a large multiplier during a period when the Kenyan economy was experiencing steady economic growth, rather than a recession; this suggests that any under-utilization of supply side capacity is not simply temporary or cyclical in rural Kenya, but may be more persistent.[49] The time frame of a cash transfer intervention could also be consequential, for instance, if the elasticity of individual labor supply or migration is larger in the long-run. Finally, the source of funding may matter: if funded through foreign aid, for example, a scaled-up national cash transfer program could also induce exchange rate effects which our data does not allow us to characterize.

Looking ahead, a traditional perspective in the case of an open economy with complete markets is that the economy should eventually revert to its previous steady-state after a

49. Recent work argues that there may be a related phenomenon of steady-state "liquidity traps" or "secular stagnation" in advanced economies as well (e.g., Rachel and Summers 2019, Mian, Straub, and Sufi 2019).

local aggregate demand shock like the one we study ends, with only transient effects on consumption and prices (Farhi and Werning 2016). However, other theoretical perspectives from international trade, economic geography, and development (e.g., Marshall 1890, Rosenstein-Rodan 1943, Murphy, Shleifer, and Vishny 1989, Krugman 1991), as well as the liquidity traps literature, suggest there could be persistent local effects of a temporary cash infusion, due to agglomeration effects, increasing returns, changes in income inequality, market structure and firm specialization, and even shifts in the social networks of traders and suppliers. Temporary cash transfers and other forms of assistance have also been shown to have effects on long-run human capital accumulation and earnings (Bouguen et al. 2019; Baird et al. 2016). An evaluation of long-run patterns of economic activity, firm dynamics, migration, and household living standards in the sample communities would provide a valuable experimental test of these theories.

References

Angelucci, Manuela, and Giacomo De Giorgi. 2009. "Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption?" *Am. Econ. Rev.* 99 (1): 486–508.

Auerbach, Alan J, Yuriy Gorodnichenko, and Daniel Murphy. 2019. *Local Fiscal Multipliers and Fiscal Spillovers in the United States,* Working Paper 25457. NBER.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-run Impacts of a Child Health Investment." *QJE* 131:1637–80.

Baird, Sarah, Craig McIntosh, and Berk Ozler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *QJE* 126 (4): 1709–53.

Banerjee, Abhijit, Rema Hanna, Gabriel E. Kreindler, and Benjamin A. Olken. 2017. "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs." *World Bank Research Observer* 32 (2): 155–184.

Bardhan, Pranab. 2005. "Theory or Empirics in Development Economics." *Economic and Political Weekly* 40 (40).

Bassi, Vittorio, Raffaela Muoio, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2019. *Achieving Scale Collectively.* Technical report.

Bastagli, F., J. Hagen-Zanker, L. Harman, V. Barca, G. Sturge, and T. Schmidt with L. Pellerano. 2016. *Cash transfers: what does the evidence say?* Overseas Development Institute.

Bouguen, A., Y. Huang, M. Kremer, and E. Miguel. 2019. "Using RCTs to Estimate Long-Run Impacts in Development Economics." *Annual Rev. of Econ.* 11:523–61.

Broda, Christian, and Jonathan A. Parker. 2014. "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption." *J. of Monetary Econ.* 68:S20–36.

Burke, Marshall, Lauren Falcao Bergquist, and E. Miguel. 2019. "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets." *QJE* 134 (2): 785–842.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Econ. and Stat.* 90 (3): 414–27.

Chodorow-Reich, Gabriel. 2019. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *AEJ: Economic Policy* 11 (2): 1–34.

Conley, Timothy G. 2008. "Spatial Econometrics." In *The New Palgrave Dictionary of Economics,* Second Edition, edited by Steven N. Durlauf and Lawrence E. Blume, 7:741–47. Houndsmills: Palgrave Macmillan.

Corbi, Raphael, Elias Papaioannou, and Paolo Surico. 2019. "Regional Transfer Multipliers." *Review of Economic Studies* 86:1901–1934.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran. 2018. "The Price Effects of Cash Versus In-Kind Transfers." *Review of Economic Studies* 86 (1): 240–81.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55.

Deaton, Angus S. 2018. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* New York: World Bank Group.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *QJE* 131 (2): 799–858.

Easterly, William. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done so Much Ill and so Little Good.* Penguin Books.

Evans, David K., and Anna Popova. 2017. "Cash Transfers and Temptation Goods." *Economic Development and Cultural Change* 65 (2): 189–221.

Farhi, Emmanuel, and Ivan Werning. 2016. "Fiscal Multipliers: Liquidity Traps and Currency Unions." In *Handbook of Macroeconomics,* edited by J.B. Taylor and H. Uhlig, 2:2417–92. Amsterdam: Elsevier.

Filmer, Deon, Jed Friedman, Eeshani Kandpal, and Junko Onishi. 2018. "Cash Transfers, Food Prices, and Nutrition Impacts on Nonbeneficiary Children." March.

Foster, Vivien, and Cecilia Briceno-Garmendia. 2010. *Africa's Infrastructure: A Time for Transformation.* Washington DC: World Bank Publications.

Haushofer, Johannes, and Jeremy Shapiro. in prep. "The Short-term Consumption Response to Unconditional Cash Transfers."

———. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *QJE* 131 (4): 1973–2042.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *QJE* 124 (4): 1403–1448.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money.* London: Macmillan.

Kraay, Aart. 2014. "Government Spending Multipliers in Developing Countries: Evidence from Lending by Official Creditors." *AEJ: Macroeconomics* 6 (4): 170–208.

Krugman, Paul. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99 (3): 483–499.

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." *The Manchester School* 22 (2): 139–191.

MacKinnon, James G., and Matthew D. Webb. 2018. "The wild bootstrap for few (treated) clusters." *Econometrics Journal* 21 (2): 114–35.

Mankiw, N. Gregory, and Matthew Weinzierl. 2011. "An Exploration of Optimal Stabilization Policy." *Brookings Papers on Economic Activity* 42 (1): 209–272.

Marshall, Alfred. 1890. *The Principles of Economics.* McMaster University.

Mian, Atif, Ludwig Straub, and Amir Sufi. 2019. *Indebted Demand.* Unpublished.

Michaillat, Pascal, and Emmanuel Saez. 2015. "Aggregate Demand, Idle Time, and Unemployment." *QJE* 130 (2): 507–569.

Miguel, Edward, and Michael Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72 (1): 159–217.

Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at Scale." *Journal of Econ. Perspectives* 31 (4): 103–24.

Murphy, Daniel. 2017. "Excess capacity in a fixed-cost economy." *European Economic Review* 91:245–60.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97 (5): 1003–1026.

Nakamura, Emi, and Jón Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104 (3): 753–92.

Parker, J.A., N.S. Souleles, D.S. Johnson, and R. McClelland. 2013. "Consumer Spending and the Economic Stimulus Payments of 2008." *AER* 103 (6): 2530–53.

Pennings, Steven. 2020. "Cross-Region Transfers in a Monetary Union: Evidence from the US and Some Implications." WB Policy Research Working Paper No. 9244.

Rachel, Lukasz, and Lawrence H. Summers. 2019. "On falling neutral real rates, fiscal policy, and the risk of secular stagnation." *Brookings Papers on Economic Activity.*

Ramey, Valerie A. 2019. "Ten Years after the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research?" *Journal of Econ. Perspectives* 33 (2): 89–114.

Rosenstein-Rodan, Paul N. 1943. "Problems of Industrialisation of Eastern and South-eastern Europe." *Economic Journal* 53:202–11.

Sadoulet, E., A. de Janvry, and Benjamin Davis. 2001. "Cash Transfer Programs with Income Multipliers: PROCAMPO in Mexico." *World Development* (6): 1043–56.

Sims, Eric, and Jonathan Wolff. 2018. "The Output and Welfare Effects of Government Spending Shocks Over the Business Cycle." *International Econ. Rev.* 59 (3): 1403–35.

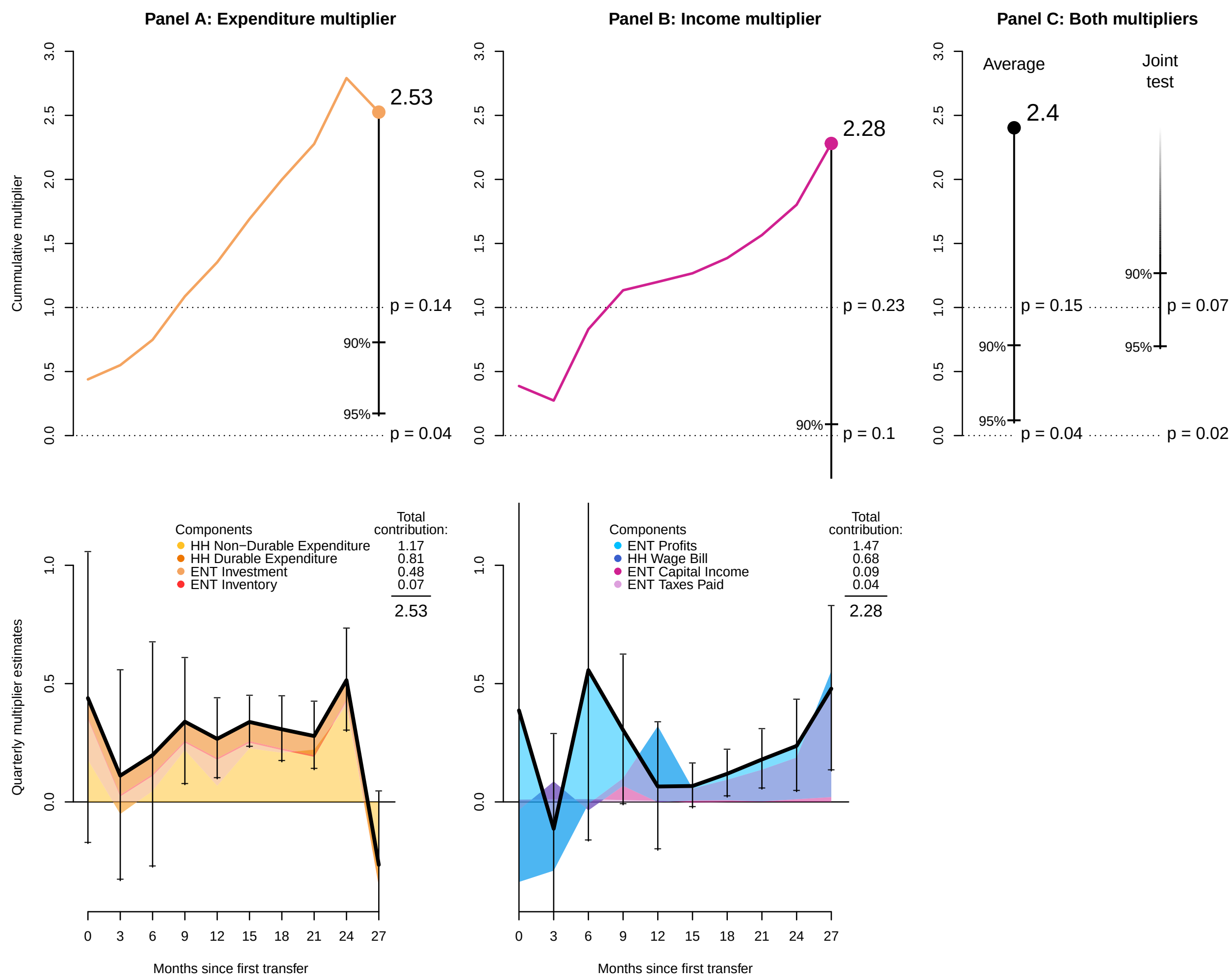
Suarez Serrato, Juan Carlos, and Philippe Wingender. 2016. *Estimating Local Fiscal Multipliers.* Unpublished.

Thome, Karen, Mateusz Filipski, Justin Kagin, J. Edward Taylor, and Benjamin Davis. 2013. "Agricultural spillover effects of cash transfers: What does LEWIE have to say?" *American Journal of Agricultural Economics* 95 (5): 1338–1344.

Walker, Michael. 2018. "Informal Taxation Responses to Cash Transfers: Experimental Evidence from Kenya." July.

World Bank. 2017. *Closing the Gap: The State of Social Safety Nets.* Washington, D.C.

Figure 1: Transfer multiplier over time



Notes: Panel A shows the cumulative expenditure multiplier over the first 29 months after start of the transfers in the top panel, and the corresponding quarterly impulse response function (IRF) in the bottom panel. The integral under this IRF yields our overall point estimate of 2.53. Colored areas below the IRF represent the different components of expenditure and the adjacent table indicates their total (over time) contribution. Darker shading indicates cases where a component turns negative in a given quarter, leading some areas to overlap. Brackets around the quarterly IRF point estimates indicate ± 1SE confidence intervals obtained from 2000 wild bootstrap replications. Whiskers below the overall point estimate indicate one-sided confidence intervals from the same bootstrap procedure, with p-values corresponding to tests of the one-sided hypotheses $H_0 : \mathbb{M} < 0$ and $H_0 : \mathbb{M} < 1$ presented at the horizontal lines at 0 and 1 respectively. Panel B repeats the same exercise for the income multiplier. Panel C presents results from aggregating the two estimators either by averaging them (left-hand side) or testing the joint null that both are less than the indicated critical values (right-hand side). In each case whiskers indicate one-sided confidence intervals obtained via the bootstrap as above.

Table 1: Expenditures, Savings and Income

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel A: Expenditure				
Household expenditure, annualized	292.98***	338.16***	333.73***	2,536.86
	(60.09)	(109.36)	(123.24)	(1,934.09)
Non-durable expenditure, annualized	186.96***	226.74**	316.62***	2,471.49
	(58.55)	(99.62)	(119.79)	(1,877.82)
Food expenditure, annualized	71.61*	133.55**	132.84**	1,578.43
	(36.93)	(63.98)	(58.58)	(1,072.31)
Temptation goods expenditure, annualized	6.51	5.88	−0.71	37.10
	(5.79)	(8.82)	(6.50)	(123.59)
Durable expenditure, annualized	95.18***	109.07***	8.41	59.44
	(12.64)	(20.23)	(12.50)	(230.90)
Panel B: Assets				
Assets (non-land, non-house), net borrowing	178.09***	182.01***	132.63*	1,132.15
	(24.61)	(44.25)	(78.32)	(1,420.22)
Housing value	372.78***	480.68***	72.58	2,033.72
	(25.25)	(38.88)	(215.70)	(5,030.37)
Land value	50.86	153.09	572.07	5,030.72
	(186.08)	(262.48)	(458.28)	(6,607.61)
Panel C: Household balance sheet				
Household income, annualized	77.62*	134.02	229.46***	1,023.45
	(43.66)	(93.83)	(88.59)	(1,634.70)
Net value of household transfers received, annualized	−1.68	−7.44	8.75	130.18
	(6.81)	(13.06)	(19.10)	(263.75)
Tax paid, annualized	1.95	−0.09	1.66	16.93
	(1.28)	(2.02)	(2.02)	(36.51)
Profits (ag & non-ag), annualized	24.70	33.73	44.08	485.20
	(23.18)	(48.95)	(45.35)	(787.10)
Wage earnings, annualized	42.51	73.72	182.99***	495.37
	(32.24)	(60.83)	(65.44)	(1,231.56)

Notes: Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households, including between 5,372 and 5,424 observations. Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households (5,448 to 5,509 observations), coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 2: Input Prices and Quantities

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	1(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel A: Labor				
Hourly wage earned by employees	0.11***	0.04	0.19*	0.70
	(0.03)	(0.04)	(0.10)	(0.89)
Household total hours worked, last 7 days	2.44	1.41	−4.70	63.20
	(1.71)	(3.69)	(3.17)	(54.14)
Panel B: Land				
Land price per acre	166.84	365.44	556.83	3,952.86
	(201.20)	(290.86)	(412.34)	(3,148.52)
Acres of land owned	−0.19	−0.10	0.08	1.42
	(0.14)	(0.09)	(0.10)	(2.37)
Panel C: Capital				
Loan-weighted interest rate, monthly	−0.01	0.01	−0.01	0.06
	(0.01)	(0.01)	(0.01)	(0.07)
Total loan amount	5.55	3.13	6.36	80.61
	(4.95)	(8.34)	(13.21)	(204.36)

Notes: Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households, including between 2,828 and 5,423 observations for variables at the household level, and 2,832 observations at the individual level for wages. Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. We have between 2,781 to 5,509 observations at the household level and 2,391 wage observations at the individual level. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. In addition, prices are quantity-weighted. That is, wages are weighted by the number of hours worked, land prices by the number of acres purchased, and interest rates by size of each loan. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 3: Enterprise Outcomes

	(1)	(2)	(3)	(4)
	Treatment Villages		**Control Villages**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Panel A: All enterprises				
Enterprise profits, annualized	10.46	67.53	32.91	323.39
	(24.35)	(41.62)	(37.27)	(691.12)
Enterprise revenue, annualized	−10.05	356.81**	244.27**	758.52
	(103.02)	(144.21)	(108.96)	(2,493.40)
Enterprise costs, annualized	−11.41	96.82**	77.02	147.73
	(29.73)	(40.92)	(48.95)	(550.11)
Enterprise wagebill, annualized	−14.72	81.69**	70.49*	120.62
	(26.48)	(33.76)	(36.52)	(492.11)
Enterprise profit margin	0.01	−0.06**	−0.06***	0.44
	(0.02)	(0.03)	(0.02)	(0.61)
Panel B: Non-agricultural enterprises				
Enterprise inventory	11.01	34.68**	16.91	192.98
	(9.14)	(14.73)	(10.80)	(504.76)
Enterprise investment, annualized	4.00	13.58	6.82	178.25
	(7.05)	(15.39)	(8.65)	(640.98)
Panel C: Village-level				
Number of enterprises	0.01	0.02	0.01	1.12
	(0.01)	(0.01)	(0.01)	(0.14)

Notes: Column 1 reports the coefficient on an indicator for treatment village, and includes an indicator for saturation status of the sublocation. Column 2 reports the total effect on enterprises in treatment villages (own-village effect plus across-village spillover) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a enterprise's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the enterprise (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer). Column 3 reports the total effect on enterprises in control villages (across-village spillover only). For each Column, we stack 3 separate regressions for own-farm enterprises, non-agricultural enterprises operated within the household, and non-agricultural enterprises operated outside the household, due to our independent sampling across these enterprise categories (Equations 10 and 11). We have between 9,997 and 10,254 observations for all enterprises, and 2,389 to 2,398 for variables we collect for non-ag enterprises only, and 653 villages. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across all enterprise categories). Each regression is weighted by inverse sampling weights and contains village-level baseline averages of the outcome variable by enterprise category when available. For monetary values, we convert effects to a per-household level by multiplying the average effect per enterprise in each enterprise category by the number of enterprises in that category, dividing by the number of households in our study area, and summing over all enterprise categories. For the number of enterprises, we run regressions at the village level, where the outcome is the number of enterprises per household in each category, we weight by the number of households in each village and sum up over all enterprise categories. For the profit margin, we weight the effects across all enterprise categories by their share in the economy, and across each enterprise by revenue, so that our estimate represents the effect on the revenue-weighted average enterprise in the economy. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 4: Output Prices

		(1)	(2)	(3)	(4)
		Overall Effects		**ATE by market access**	
		ATE	Average maximum effect (AME)	below median	above median
All goods		0.0010*	0.0042	0.0017*	0.0007
		(0.0006)	(0.0031)	(0.0009)	(0.0007)
By tradability	More tradable	0.0014	0.0062	0.0023	0.0021
		(0.0015)	(0.0082)	(0.0023)	(0.0018)
	Less tradable	0.0009	0.0034	0.0015	0.0001
		(0.0006)	(0.0032)	(0.0011)	(0.0008)
By sector	Food items	0.0009	0.0036	0.0016	0.0002
		(0.0006)	(0.0033)	(0.0012)	(0.0008)
	Non-durables	0.0014	0.0061	0.0026	0.0019
		(0.0017)	(0.0089)	(0.0026)	(0.0019)
	Durables	0.0019*	0.0070	−0.0009	0.0034**
		(0.0011)	(0.0061)	(0.0011)	(0.0016)
	Livestock	−0.0008	−0.0027	−0.0008*	−0.0017
		(0.0010)	(0.0052)	(0.0004)	(0.0020)
	Temptation goods	−0.0011	−0.0112	−0.0008	−0.0003
		(0.0026)	(0.0143)	(0.0036)	(0.0035)

Notes: Each row represents a regression of the logarithm of a price index on the "optimal" number of lags and distance buffers of per capita Give Directly transfers in each buffer. Price indices are based on 311,138 non-missing price quotes for 70 commodities and products. For each product, we take the logarithm of the median price quote in a market-month, and create our market price indices as an expenditure weighted average of these median price quotes across all goods in that market-month. Regressions include a panel of 1,734 market-by-month observations. The number of radii bands and lags is chosen sequentially by minimizing the BIC, as pre-specified, for the overall price index, which selects a 4km radius; subcomponents use this value as well. Regressions include a full set of market and month fixed effects. Column 1 reports the implied ATE, calculated by evaluating the "optimal" regression specification at the average level of treatment intensity between September 2014 and March 2017, the time during which transfers went out. Column 2 reports the average maximum effect, calculated at the average across all markets of the month in which the largest per capita transfers went into a market's neighborhood (up to the largest buffer selected by the algorithm). Columns 3 and 4 break down the ATE by market access, defined as $MA_m = \sum_{r=0}^{10} r^{-\theta} N_r$, where $\theta = 8$ and N_r is the population in in the $r-2$ to r km buffer around each market. Standard errors (in parentheses) are as in and we allow for spatial correlation up to 10km and autocorrelation up to 12 months. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table 5: Transfer Multiplier Estimates

	(1) $\mathbb{M}$ Estimate	(2) H_0: $\mathbb{M} < 0$ p-value	(3) H_0: $\mathbb{M} < 1$ p-value
Panel A: Expenditure multiplier	2.53	0.04**	0.14
	(1.42)		
Household non-durable expenditure	1.17	0.19	
	(1.32)		
Household durable expenditure	0.81	0.00***	
	(0.05)		
Enterprise investment	0.48	0.13	
	(0.42)		
Enterprise inventory	0.07	0.02**	
	(0.03)		
Panel B: Income multiplier	2.28	0.10*	0.23
	(1.73)		
Enterprise profits	1.47	0.13	
	(1.28)		
Household wage bill	0.68	0.27	
	(1.15)		
Enterprise capital income	0.09	0.31	
	(0.17)		
Enterprise taxes paid	0.04	0.08*	
	(0.03)		
Panel C: Expenditure and income multipliers			
Average of both multipliers	2.40	0.04**	0.15
	(1.38)		
Joint test of both multipliers		0.02**	0.07*

Notes: Results are from the joint estimation of expenditure and income multipliers. Column 1 reports point estimates of both multipliers and their respective components. Each component is estimated individually and the multiplier is obtained by aggregating components as described in the main text. Effects of the cash infusion on flow variables (non-durable consumption, investment, wages, profits, capital income, and taxes) are obtained by dynamically estimating effect sizes over 29 months after the first transfer and computing the integral under this curve (Equation 7). Effects on remaining stock variables are the estimated the total endline treatment effects (Equations 2, 3 and 11). Transfer amounts and outcome variables are deflated to January 2015 using the overall consumer price index in the geographically closest market. Standard errors are computed from 2,000 replications of a wild clustered bootstrap, which re-allocates within-sublocation Rademacher-perturbed residuals from the main population regressions to fitted outcome values to create perturbed samples. Columns 2 and 3 conduct one-sided tests of each multiplier estimate $\mathbb{M}$ against 0 and 1 respectively, using the bootstrapped distributions of $\mathbb{M}$. Panel C conducts two tests regarding both multipliers. The first row computes the average of both estimates and conducts tests on this average using the same bootstrap procedure. The last row reports p-values from joint tests of both multipliers against the same nulls. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

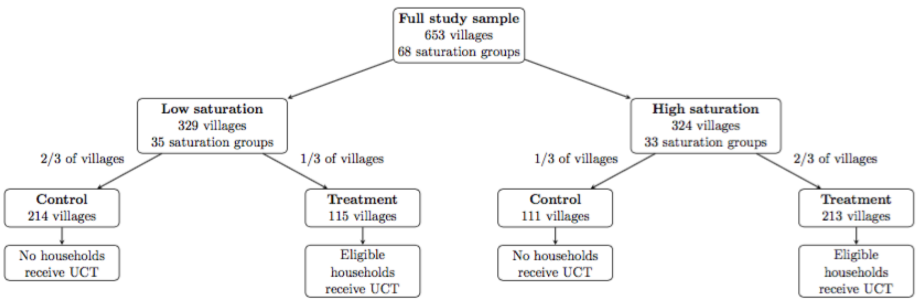
Appendix

Table of Contents

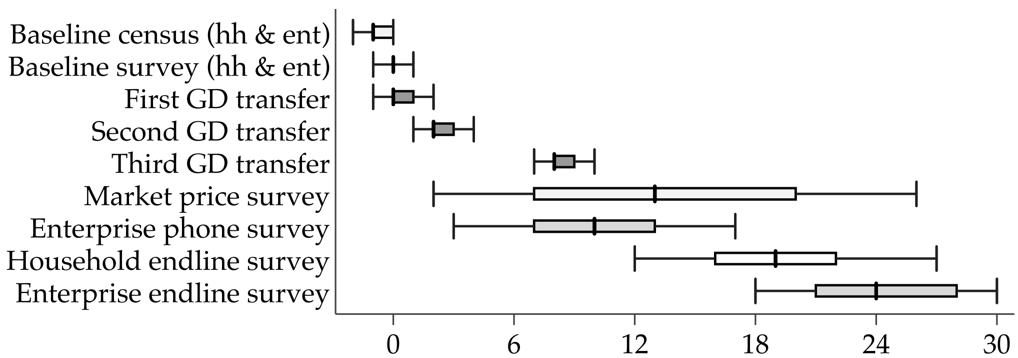
A Supporting figures & tables

Figure A.1: Study design and timeline

(a) Randomization

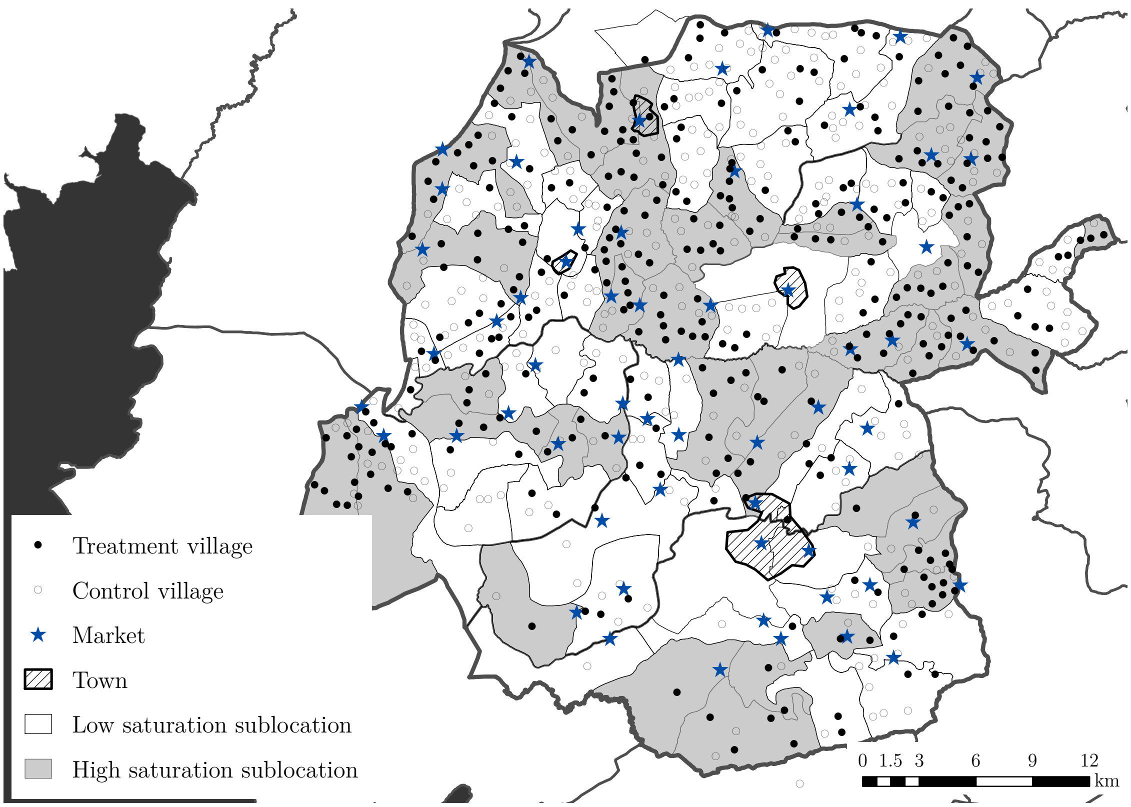


(b) Timing relative to experimental start



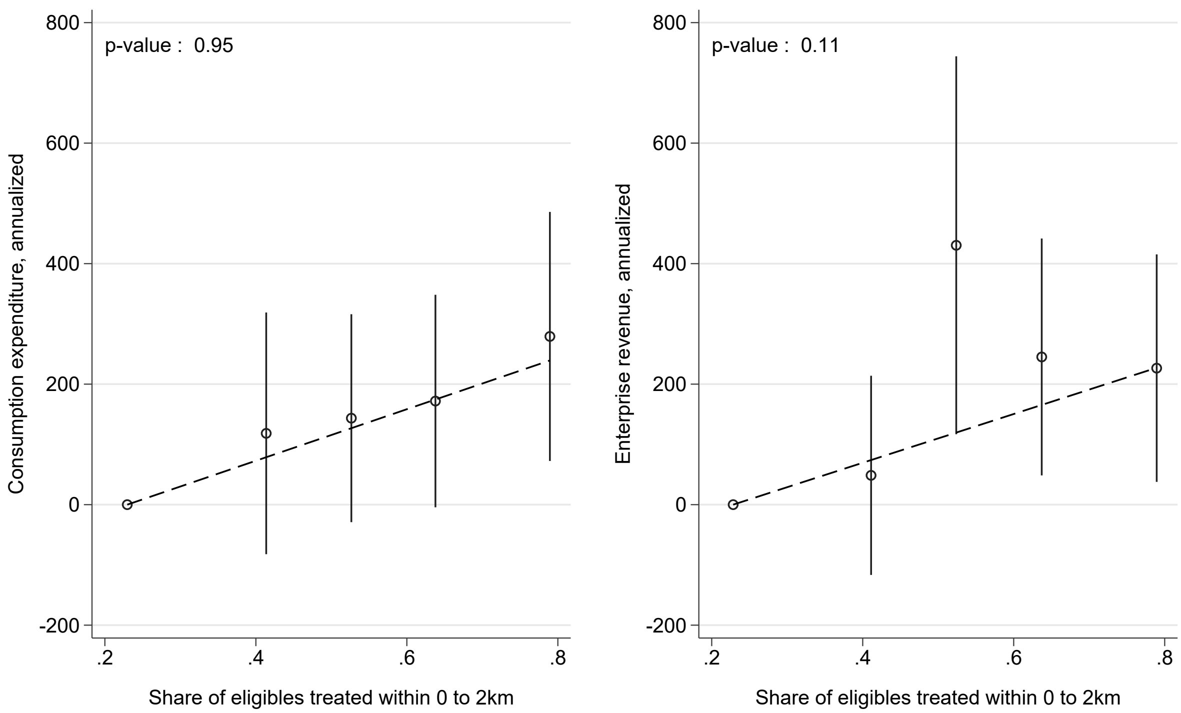
Notes: Panel A illustrates the two-level randomized controlled trial experimental design. 653 villages were grouped into 68 saturation groups based on the sublocation (the administrative unit directly above the village level) in which they are located. Saturation groups were then randomly assigned to either high or low saturation status. In the 33 high saturation groups, two-thirds of villages were assigned to treatment status, while in the 35 low saturation groups, one-third of villages are assigned to treatment status. In the 328 treatment villages, all eligible households received an unconditional cash transfer, while no households within control villages received a transfer. Panel B plots the 5th, 25th, 50th, 75th and 95th percentiles of study activities. Timing is reported relative to the anticipated start of activities in each village (the "experimental start"). The experimental start for a village is calculated based on the random ordering of treatment and control villages that both GD and research team field enumerators worked in, as well as GD's mean monthly pace of enrolling villages in the subcounty in which the village is located. As markets were not assigned to treatment, we use the first date transfers were distributed within the subcounty in which the market is located. The value of the first GD transfer is USD 151 PPP, while the second and third are both USD 860 PPP.

Figure A.2: Study area



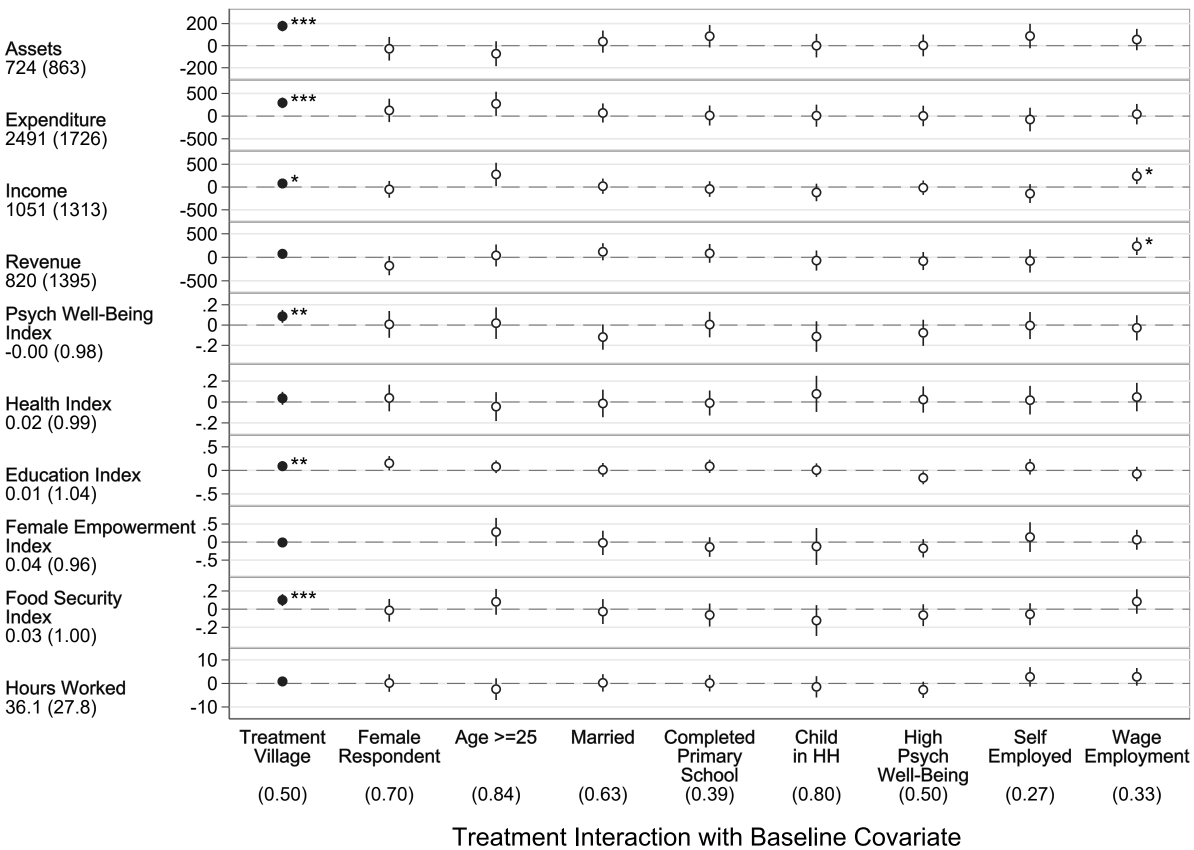
Notes: This figure plots study villages, sublocation boundaries, and weekly markets in the study area in Siaya County, Kenya. Control villages are denoted by hollow circles, treatment villages are denoted by solid circles, and blue stars indicate the locations of markets. High saturation sublocations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

Figure A.3: Non-linear Spillover Estimates



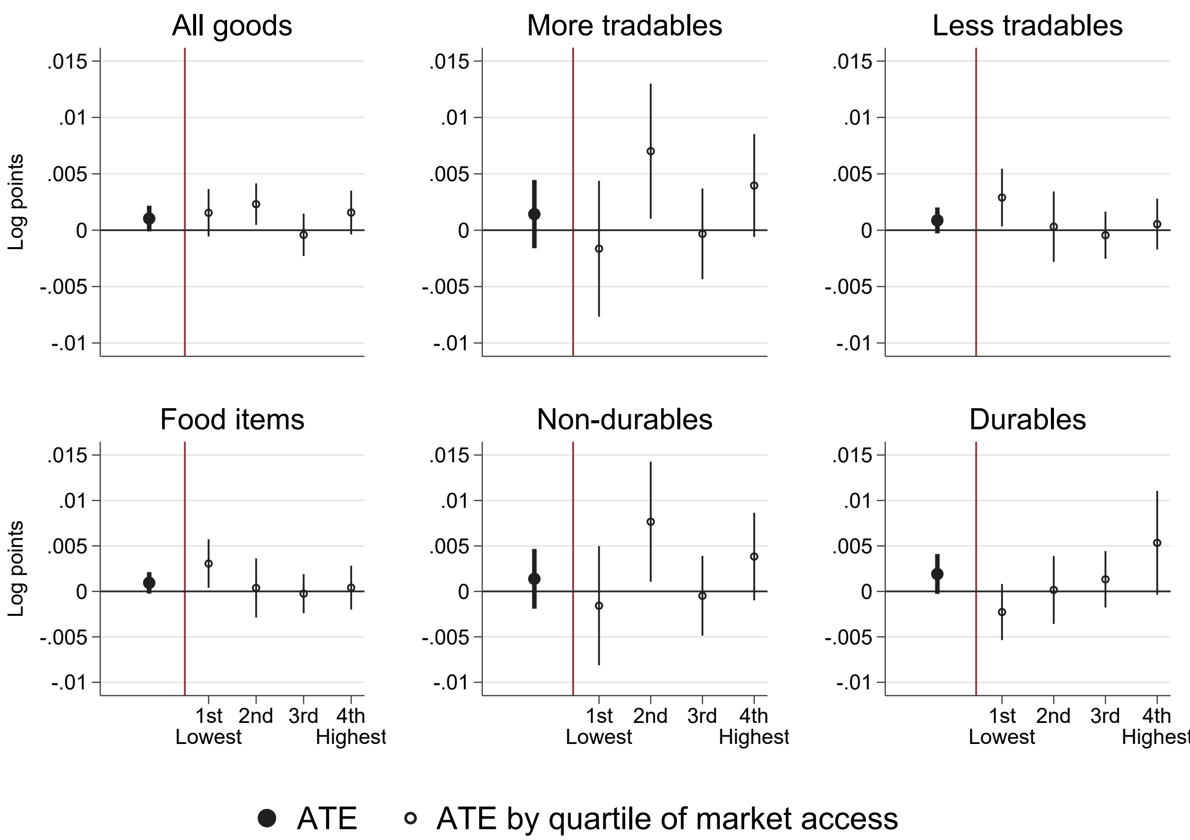
Notes: Each panel represents a reduced-form regression of household consumption and enterprise revenue on 5 quintile bins of the share of eligibles assigned to treatment 2km around each household / enterprise. Effects are relative to the lowest quantile. For consumption, we control for direct effects by including eligibility and a dummy for treatment status of each household. For enterprises, we include interactions with dummies for 3 enterprise types (within homestead, outside homestead, and own-farm). We then sum and weight coefficients to obtain total revenue effects per household in our study area. We control for baseline revenue at the village-by-enterprise type level and use inverse sampling weights. 95% confidence intervals are obtained using Conley (2008). Dashed lines start at zero, the slope coming from the same regression, with quantile bins of treatment intensity replaced by a linear term (and weighted across enterprise types as above). We cannot formally reject that our estimated non-linear regression is linear, i.e. that $\frac{\beta_2-\beta_1}{\Delta X_2} = ... = \frac{\beta_n - beta_{n-1}}{\Delta X_n}$. The p-values of this test are 0.95 and 0.11 for consumption and revenue respectively. We did the same test for all 10 pre-specified primary outcomes and treated / untreated households separately; we cannot reject linearity at the 10% level for any of them.

Figure A.4: Little heterogeneity in pre-specified primary outcomes



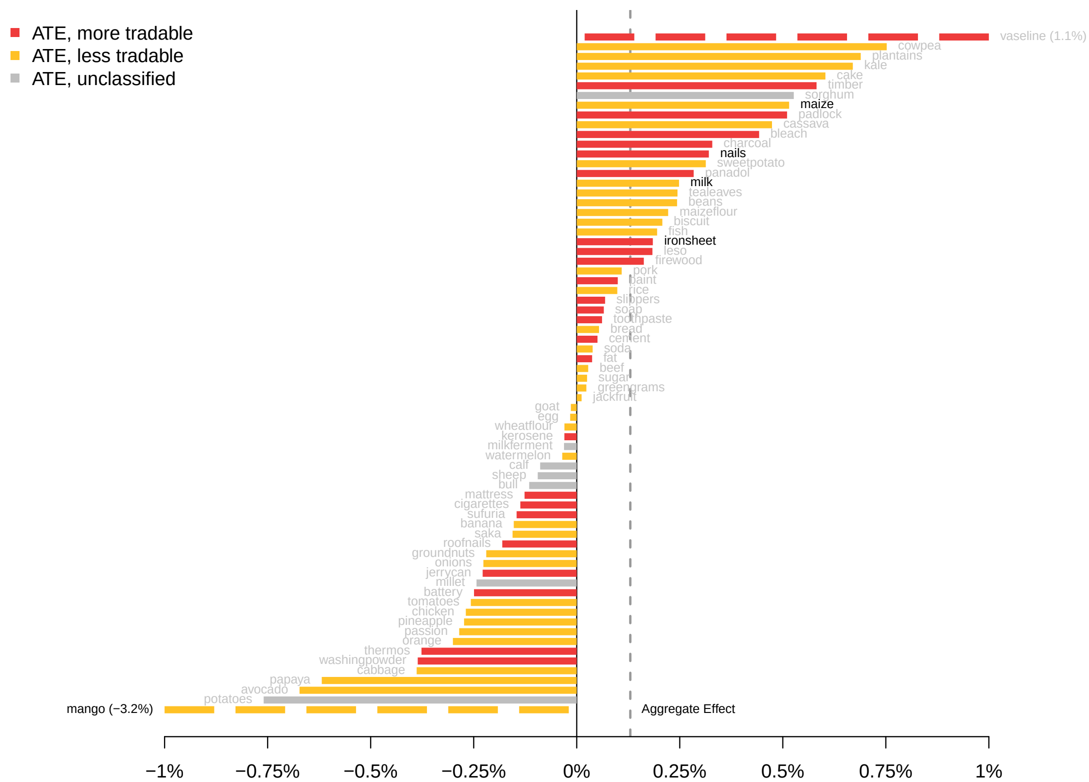
Notes: This figure presents estimates for treatment effect heterogeneity for eligible households in pre-specified primary outcomes along 8 pre-specified dimensions of heterogeneity (Haushofer et al. 2017a). Each plotted coefficient is from a separate regression. Each row represents a separate primary outcome; the mean (SD) for eligible households in control, low saturation villages is reported below the outcome label. The first column (Treatment Village) plots estimated effects for the coefficient on an indicator for being in a treatment village from Equation (1), where the sample is restricted to eligible households. Columns 2 through 8 plot the coefficient on the interaction term of the listed baseline covariate with the treatment village indicator; this interaction term and baseline covariate are added to Equation (1). Values in parentheses on the x-axis denote the mean of the baseline covariate. Standard errors are clustered at the village level. Reported significance levels correspond to FDR q-values, calculated following Benjamini, Krieger, and Yekutieli (2006). * denotes significance at 10 pct., ** denotes significance at 5 pct., and *** denotes significance at 1 pct. level.

Figure A.5: Output price effects by market access



Notes: Each panel represents a regression of the logarithm of a price index on the "optimal" number of lags and distance buffers of per capita GiveDirectly transfers in each buffer, as calculated for the overall price index. The number of radii bands and lags is chosen sequentially by minimizing the BIC, as pre-specified, for the overall price index. Regressions include a full set of market and month fixed effects. We report the implied ATE, calculated by evaluating the "optimal" regression specification at the average level of treatment intensity between September 2014 and March 2017, the time during which transfers went out. Bold markers correspond to the ATE across all markets. Hollow markers break down this average by quartiles of market access (with low market access referring to more remote markets), defined as $MA_m = \sum_{r=1}^{10} r^{-\theta} N_r$, where $\theta = 8$ and N_r is the population in in the $r-1$ to r km buffer around each market. Bars represent 95% confidence intervals based on standard errors as in Conley (2008), where we allow for spatial correlation up to 10km and autocorrelation up to 12 months.

Figure A.6: Output price effects at the product level



Notes: Each bar represents a regression of the logarithm of a median price index for each good, using a 4km distance buffer and no lags (the "optimal" number of lags and distance buffers of per capita GiveDirectly transfers for the overall price index). The number of radii bands and lags is chosen sequentially by minimizing the BIC, as pre-specified. Regressions include a full set of market and month fixed effects. Colors denote our classification into more tradable vs. less tradable goods. For each good, we report the implied ATE, 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. Bold product names indicate significance at the 95% level.

Table A.1: Household Assets by Productivity Status

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Assets (non-land, non-house), net borrowing	178.09***	182.01***	132.63*	1,132.15
	(24.61)	(44.25)	(78.32)	(1,420.22)
Productive Agricultural Assets	4.26***	4.16**	−0.38	32.51
	(0.93)	(1.96)	(2.47)	(38.94)
Potentially Productive Assets	89.97***	52.78	35.86	700.68
	(25.86)	(49.31)	(65.83)	(1,025.41)
Livestock Assets	50.61***	44.82	−7.01	462.24
	(17.04)	(27.90)	(35.76)	(723.43)
Non-Ag Assets	37.03***	24.60	25.63	219.04
	(10.43)	(22.85)	(23.15)	(424.05)
Non-Productive Assets	78.81***	92.59***	52.33*	449.30
	(9.28)	(14.28)	(29.60)	(468.64)

Notes: This table presents results on household asset ownership based on classifications of assets by productivity status. Productive agricultural assets include agricultural tools. Potentially productive assets include livestock and non-agricultural assets, made up of the following: bicycle, motorcycle, car, boat, kerosene stove, sewing machine electric iron, computer, mobile phone, car battery, solar (panels or system), and generators. Non-productive assets include: radio/cd player, kerosene lantern, bed, mattress, bednet, table, sofa, chair, cupboards, clock, television, iron sheets. Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2). For this analysis, the sample is restricted to eligible households, including 5,420 observations. Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3). We have 5,505 observations. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC for total assets (Row 1). Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.2: Enterprise revenue effects by sector

	(1)	(2)	(3)	(4)
	Treatment Villages		**Control Villages**	
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Retail revenue	70.95*	140.92**	65.14	236.20
	(42.22)	(70.75)	(57.15)	(415.34)
Manufacturing revenue	−30.40	109.34**	109.58**	81.27
	(50.46)	(51.97)	(55.57)	(177.27)
Services revenue	−77.63*	17.77	52.70*	115.19
	(40.10)	(43.30)	(28.93)	(175.93)
Agriculture revenue	20.64	41.78*	16.38	308.18
	(13.91)	(23.95)	(19.83)	(365.39)

Notes: Column 1 reports the coefficient on an indicator for treatment village, and includes an indicator for saturation status of the sublocation (Equation 10). Column 2 reports the total effect on enterprises in treatment villages (own-village effect plus across-village spillover) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a enterprise's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the enterprise (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer). Column 3 reports the total effect on enterprises in control villages (across-village spillover only). For non-agricultural sectors (retail, services and manufacturing), we stack 2 separate regressions for non-agricultural enterprises operated within the household, and non-agricultural enterprises operated outside the household, due to our independent sampling across these enterprise categories (as in Equation 11). We have 1,300 observations for retail enterprises, 576 for manufacturing, 400 for services and 7,896 for agriculture. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across all enterprise categories). Each regression is weighted by inverse sampling weights and contains village-level baseline averages of the outcome variable by enterprise category when available. We convert effects to a per-household level by multiplying the average effect per enterprise in each enterprise category by the number of enterprises in that category, dividing by the number of households in our study area, and summing over all enterprise categories. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.3: Enterprise outcomes by owner eligibility

	(1)	(2)	(3)	(4)
	Recipient Owners		**Non-Recipient Owners**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Panel A: All enterprises				
Enterprise profits, annualized	12.03	29.36	49.61**	323.39
	(13.44)	(22.01)	(24.30)	(691.12)
Enterprise revenue, annualized	71.90***	148.44	197.78***	758.52
	(26.40)	(97.10)	(51.67)	(2,493.40)
Enterprise costs, annualized	29.55***	46.83***	45.82**	147.73
	(9.98)	(15.41)	(18.37)	(550.11)
Enterprise wagebill, annualized	25.29***	43.83***	43.63**	120.62
	(9.17)	(14.40)	(17.73)	(492.11)
Enterprise profit margin	−0.04**	−0.07**	−0.03	0.44
	(0.02)	(0.03)	(0.03)	(0.61)
Panel B: Non-agricultural enterprises				
Enterprise inventory	2.88	7.74	5.58	192.98
	(2.79)	(7.33)	(5.56)	(504.76)
Enterprise investment, annualized	−5.15	−15.61	5.50	178.25
	(5.35)	(16.17)	(8.25)	(640.98)
Panel C: Village-level				
Number of enterprises	0.00	0.03***	0.00	1.12
	(0.01)	(0.01)	(0.01)	(0.14)

Notes: Column 1 reports the coefficient on an indicator for treatment village, and includes an indicator for saturation status of the sublocation, among matched eligible enterprise owners (Equation 10). Column 2 reports the total effect on enterprises with a treated owner relative to eligible owners in control villages (own-village effect plus across-village spillover) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a enterprise's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the enterprise (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer). We have between 5,424 and 5,555 owner-matched observations for all enterprises, and 415 for non-ag outcomes. Column 3 reports the total effect on enterprises with untreated owners (spillover only), where we have between 6,584 to 6,739 observations for all enterprises, and 1,454 to 1,459 for non-ag outcomes. For each column, we stack 3 separate regressions for own-farm enterprises, non-agricultural enterprises operated within the household, and non-agricultural enterprises operated outside the household, due to our independent sampling across these enterprise categories (Equation 11. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across all enterprise categories). Each regression is weighted by inverse sampling weights and contains village-level baseline averages of the outcome variable by enterprise category when available. We convert effects to a per-household level by multiplying the average effect in each enterprise category by the number of enterprises in that category, dividing by the number of households in our study area, and summing over all enterprise categories. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.4: Input prices and quantities: additional labor supply outcomes

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel A: Quantities				
Household hours worked on own farm	2.07*	0.98	−6.28**	35.33
	(1.15)	(2.30)	(2.61)	(38.81)
Individual hours worked in self-employment	1.80	4.23**	−1.37	26.80
	(1.14)	(1.96)	(1.76)	(23.54)
Individual hours employed last week	0.52	−1.38	2.49	23.62
	(0.98)	(2.32)	(2.67)	(25.96)
Individual hours employed last week in agriculture	−1.54***	−2.28***	0.33	6.01
	(0.56)	(0.75)	(1.11)	(12.78)
Individual hours employed last week not in agriculture	1.67	0.62	1.91	17.09
	(1.03)	(2.31)	(2.65)	(26.41)
Panel B: Prices				
Hourly wage earned by employees	0.11***	0.04	0.19*	0.70
	(0.03)	(0.04)	(0.10)	(0.89)
Hourly wage earned by employees in agriculture	0.15**	0.21**	−0.06	0.67
	(0.06)	(0.08)	(0.13)	(0.67)
Hourly wage earned by employees not in agriculture	0.04	0.08	0.20	1.09
	(0.08)	(0.10)	(0.23)	(1.45)

Notes: Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households, including between 5,420 observations at the household level, and between 1,201 and 4,085 observations for individual-level outcomes. Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. We have between 5,505 household observations, and between 1,019 and 3,486 individuals. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. In addition, prices are quantity-weighted. That is, wages are weighted by the number of hours worked. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.5: Input prices and quantities: additional land outcomes

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel A: Quantities				
Acres of land owned	−0.19	−0.10	0.08	1.42
	(0.14)	(0.09)	(0.10)	(2.37)
Acres of land rented out	−0.04	−0.05	0.06	0.93
	(0.11)	(0.21)	(0.18)	(0.91)
Acres of land rented in	0.03	0.04	0.08	0.70
	(0.03)	(0.06)	(0.07)	(0.64)
Acres of land used for crops	0.03	−0.03	0.09	0.96
	(0.02)	(0.04)	(0.06)	(1.18)
Panel B: Prices				
Land price per acre	166.84	365.44	556.83	3,952.86
	(201.20)	(290.86)	(412.34)	(3,148.52)
Monthly land rental price per acre	−0.05	−0.02	1.80	9.71
	(0.56)	(0.96)	(1.41)	(8.33)
Total ag land rental costs	6.95***	8.97*	10.14	51.76
	(2.47)	(5.21)	(9.39)	(39.67)

Notes: Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households, including between 352 and 5,418 observations (indicating land markets are often thin). Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. We have between 348 and 5,505 observations. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. In addition, prices are quantity-weighted. That is, land prices and rental rates are weighted by land size. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.6: Non-market Outcomes and Externalities

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Psychological well-being index	0.09***	0.12*	0.08	0.01
	(0.03)	(0.07)	(0.06)	(1.01)
Health index	0.04	0.06	0.01	0.03
	(0.03)	(0.06)	(0.05)	(1.01)
Food security index	0.10***	0.05	0.08	0.01
	(0.03)	(0.07)	(0.06)	(1.00)
Children food security	0.13***	0.17**	0.09	−0.04
	(0.04)	(0.08)	(0.09)	(1.12)
Education index	0.09**	0.09*	0.10*	0.01
	(0.04)	(0.05)	(0.06)	(1.02)
Female empowerment index	−0.01	0.08	0.09	0.05
	(0.07)	(0.14)	(0.15)	(0.94)
Security index	0.11***	−0.02	−0.02	0.03
	(0.04)	(0.07)	(0.07)	(0.96)

Notes: Outcome indices in each row are calculated as weighted, standardized indices of multiple survey questions, as described in detail in Appendix C.1. Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households, including between 4,121 and and 5,423 observations (and a subset of 1,118 for female empowerment). Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. We have between 4,048 and 5,509 observations (and a subset of 978 for female empowerment). The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. The number of radii bands included in Columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. Standard errors are clustered at the sublocation in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.7: Inequality

	(1)	(2)	(3)	(4)
	Treatment Villages		**Control Villages**	
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Panel A: Expenditure				
Gini coefficient	0.7	0.8	0.2	32.4
	(0.7)	(1.3)	(1.1)	(7.8)
Counterfactual Gini coefficient	−1.1*	−2.2	0	32.4
	(0.7)	(1.3)		(7.8)
P-value: effect = counterfactual effect	p=0.08	p=0.05	p=0.84	
Panel B: Assets				
Gini coefficient	−1.1	2.1	2.8**	45.4
	(0.9)	(1.6)	(1.4)	(10.1)
Counterfactual Gini coefficient	−7.5***	−6.9***	0	45.5
	(0.8)	(0.5)		(10.8)
P-value: effect = counterfactual effect	p=0.00	p=0.00	p=0.04	

Notes: This table reports results on village level inequality as measured by Gini coefficients (0-100). Panel A presents expenditure-based Gini coefficients and Panel B presents assets-based Gini coefficients. For each panel, the first row presents results on actual Gini coefficients measured from our data. The second row estimates the same specifications as the first row, but using counterfactual Gini coefficients assuming that only recipient households gained from the cash transfers, and untreated households experienced no spillovers. We construct a hypothetical consumption expenditure and assets distribution from its baseline distribution (for assets) or by imputing a baseline distribution based on endline non-missing values in control and low-saturation villages (for expenditure). We add in the associated gain, assuming recipients spend 66% of the transfer on consumption, and 34% on assets, following the relative magnitude of the point estimates on expenditure and assets in Table 1. This is also in line with our preferred dynamic MPC estimates, where we find recipients spent 93% of the transfer in the first 29 months, 63% on non-durables and 30% on durable assets (see Appendix 4.2 for details). The p-value reported in the third row tests if the actual effect (Row 1) equals the counterfactual effect (Row 2). Gini estimates and effect estimates are weighted by inverse sampling probabilities and village size. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 and 3. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.8: Expenditures, Savings and Income: Extended version

	Recipient households		Non-recipient households			
	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control Eligibles	Ineligibles	Control, low-saturation mean (SD)
Panel A: Expenditure						
Household expenditure, annualized	292.98***	338.16***	333.73***	21.11	410.25***	2,536.86
	(60.09)	(109.36)	(123.24)	(83.76)	(147.86)	(1,934.09)
Non-durable expenditure, annualized	186.96***	226.74**	316.62***	24.77	388.06***	2,471.49
	(58.55)	(99.62)	(119.79)	(79.04)	(144.91)	(1,877.82)
Food expenditure, annualized	71.61*	133.55**	132.84**	10.64	162.75**	1,578.43
	(36.93)	(63.98)	(58.58)	(50.09)	(71.29)	(1,072.31)
Temptation goods expenditure, annualized	6.51	5.88	−0.71	10.65	−3.49	37.10
	(5.79)	(8.82)	(6.50)	(8.02)	(7.80)	(123.59)
Durable expenditure, annualized	95.18***	109.07***	8.41	5.68	9.07	59.44
	(12.64)	(20.23)	(12.50)	(16.83)	(14.99)	(230.90)
Panel B: Assets						
Assets (non-land, non-house), net borrowing	178.09***	182.01***	132.63*	−12.39	168.13*	1,132.15
	(24.61)	(44.25)	(78.32)	(39.90)	(98.03)	(1,420.22)
Housing value	372.78***	480.68***	72.58	27.18	83.69	2,033.72
	(25.25)	(38.88)	(215.70)	(37.31)	(268.14)	(5,030.37)
Land value	50.86	153.09	572.07	186.11	666.54	5,030.72
	(186.08)	(262.48)	(458.28)	(290.34)	(543.51)	(6,607.61)
Panel C: Household balance sheet						
Household income, annualized	77.62*	134.02	229.46***	81.25	265.74**	1,023.45
	(43.66)	(93.83)	(88.59)	(59.37)	(108.12)	(1,634.70)
Net value of household transfers received, annualized	−1.68	−7.44	8.75	−6.83	12.56	130.18
	(6.81)	(13.06)	(19.10)	(10.27)	(23.18)	(263.75)
Tax paid, annualized	1.95	−0.09	1.66	−0.92	2.29	16.93
	(1.28)	(2.02)	(2.02)	(1.65)	(2.39)	(36.51)
Profits (ag & non-ag), annualized	24.70	33.73	44.08	−6.23	56.39	485.20
	(23.18)	(48.95)	(45.35)	(37.99)	(56.34)	(787.10)
Wage earnings, annualized	42.51	73.72	182.99***	90.02**	205.75**	495.37
	(32.24)	(60.83)	(65.44)	(39.12)	(80.09)	(1,231.56)

Notes: See Table 1 for a description of Columns 1 to 3 and 6. Columns 4 and 5 break out the total effects from Column 3 separately for eligible households in control villages and ineligible households (in both treatment and control villages), respectively. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 - 5. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table A.9: Expenditures, savings and income results excluding respondents that migrated

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel 1: Full Sample				
Respondent migrated	0.01	−0.03	0.00	0.05
	(0.01)	(0.02)	(0.01)	(0.21)
Net change in household members since baseline	0.03	0.02	−0.18**	−0.10
	(0.04)	(0.07)	(0.08)	(1.30)
Household size	0.02	0.03	−0.10	4.05
	(0.05)	(0.08)	(0.08)	(2.35)
Panel 2: Non-Migrant Sample				
Panel 2.A: Expenditure				
Household expenditure, annualized	311.86***	376.53***	325.42***	2,511.93
	(61.09)	(109.97)	(124.71)	(1,927.10)
Non-durable expenditure, annualized	200.13***	260.62***	307.06**	2,445.76
	(59.08)	(99.57)	(121.09)	(1,869.34)
Food expenditure, annualized	80.71**	152.03**	123.95**	1,572.86
	(37.98)	(67.87)	(57.68)	(1,069.81)
Temptation goods expenditure, annualized	4.45	1.89	−0.63	37.93
	(6.15)	(8.97)	(6.71)	(125.55)
Durable expenditure, annualized	102.17***	113.42***	8.54	60.04
	(13.19)	(21.60)	(12.85)	(231.72)
Panel 2.B: Assets				
Assets (non-land, non-house), net borrowing	174.46***	172.13***	136.47*	1,145.65
	(25.24)	(47.36)	(80.18)	(1,414.90)
Housing value	399.83***	478.90***	37.51	2,097.78
	(26.38)	(39.92)	(222.81)	(5,133.46)
Land value	52.15	83.35	559.27	5,141.46
	(193.27)	(286.26)	(471.63)	(6,687.63)
Panel 2.C: Household balance sheet				
Household income, annualized	37.82	84.56	201.11**	993.03
	(43.18)	(93.36)	(91.13)	(1,600.50)
Net value of household transfers received, annualized	0.68	−10.88	10.52	135.91
	(7.03)	(14.56)	(19.75)	(266.53)
Tax paid, annualized	1.59	−0.94	1.41	16.66
	(1.32)	(2.13)	(2.00)	(35.72)
Profits (ag & non-ag), annualized	11.55	−2.98	23.14	488.95
	(23.08)	(51.93)	(45.16)	(786.46)
Wage earnings, annualized	16.07	67.20	175.65***	461.17
	(31.58)	(58.98)	(65.43)	(1,185.26)

Notes: Panel 1 presents estimates of migration impacts on 3 indicators of migration: Whether the respondent themselves migrated out of the study area, the net change in household members since baseline, and the endline household size. Panel 2 reports results from Table 1 for respondents that have not migrated, where migration is defined as living in another administrative sublocation for over 4 months. See Table 1 for a descriptions of Columns 1-4. In Panel A, we have between 5,403 and 5,422 observations for columns 1-2 and 5,489 and 5,508 for column 3. In Panels B and C, we have 4,982 to 5,024 observations in columns 1-2 and 5,170 to 5,220 observations in column 3. Standard errors are clustered at the village level in Column 1, and calculated following Conley (2008) using a uniform kernel out to 10 km in Columns 2 - 5. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

B Details on study design and intervention

B.1 Cash transfer intervention

The NGO GiveDirectly (GD) provides cash transfers to poor households, and for the purposes of this study, enrolled households with grass-thatched roofs in villages assigned to treatment. GD worked on a rolling basis across villages in the study area. The village order was randomized. GD's enrollment process in treatment villages consisted of the following 6 steps:

1. Village meeting (*baraza*): Before beginning work in a village, GD held a meeting for all households in the village to inform residents that GD would be working in their village and explain the program and GD as an organization. To prevent gaming, the eligibility criteria were not disclosed.

2. Census: GD staff then conducted a household census of the village, collecting names of household members, contact information, and information about housing materials. The information on housing materials was used to determine program eligibility.

3. Registration: Households identified as eligible based on the household census were visited by GD's registration team. During these visits, GD staff confirmed the eligibility of the household, informed the household of their eligibility, and registered the household for the program. Households could select the member that they wished to be registered for the program. This visit was the point at which households learned they would be receiving transfers, as well as the amount of the transfers, the transfer schedule, and the fact that the transfer was unconditional.[50] Households were instructed and coached on how to register for M-Pesa, which was a prerequisite for being able to receive transfers. Households that did not have a mobile phone were given the option to purchase one from GD staff, the cost of which was deducted from the transfer amount.

4. Back-check: All registered households were back-checked to confirm eligibility in advance of transfers being sent. Importantly, the census, registration, and back-check teams consisted of separate staff members; this fact, and the multiple eligibility confirmations, were security measures to prevent gaming by households and field staff.

5. Transfers: Transfers were made in a series of three payments via M-Pesa, according to the following schedule: (i) A token transfer of KES 7,000 (USD 151 PPP) was sent once a majority of eligible households within the village had completed their backchecks, to ensure that the system was working properly, to ensure that the system was working properly. (ii) Two months after the token transfer, a first large installment of KES 40,000 (USD 860 PPP) was sent. (iii) Six months later (eight months after the token transfer), a second and final large installment of KES 40,000 was sent. If households elected to receive a mobile phone from GD, this cost (KES 1600 or USD 34 PPP) was subtracted from the second large installment. Transfers were typically sent at a

50. To emphasize the unconditional nature of the transfer, households were provided a brochure that listed a large number of potential uses of the transfer.

single time per month (usually around the 15th) to all households scheduled to receive transfers.

6. Follow-up: After transfers were sent, GD staff followed up by phone with transfer recipients to ensure that transfers were received. In addition, recipients could contact a GD helpline with questions. If GD staff learned that household conflicts had arisen as a result of the transfers, transfers were occasionally delayed while these problems were worked out.

B.2 Randomization details

Villages were randomly assigned to treatment status following the two-level randomization design described in Figure A.1a. The randomization was conducted in two batches as GD expanded its operations, with the first batch covering villages in Alego subcounty, and the second batch covering villages in Ugunja and Ugenya subcounties.

In Alego, we compiled a list of rural villages eligible for GD expansion. We then grouped sublocations into 23 saturation groups, ensuring that each saturation group was formed from contiguous sublocations, had at least three study villages, and (where possible) the number of study villages was a multiple of three (given that either one-third or two-thirds of villages are assigned to treatment within each sublocation). In 11 sublocations, we declared the sublocation itself as the saturation group. The remaining 13 saturation groups were formed by combining contiguous sublocations into saturation groups. In this manner, the 39 sublocations in Alego were allocated to 23 saturation groups, which were later randomized into high- and low-saturation status.

GD had worked in 193 villages in Alego prior to the start of this study. To account for previous participation in GD's program, we stratified assignment of high and low saturation by the level of previous exposure to the GD program within the saturation group, measured as the share of villages covered by a previous GD campaign, splitting the exposure level at the median.

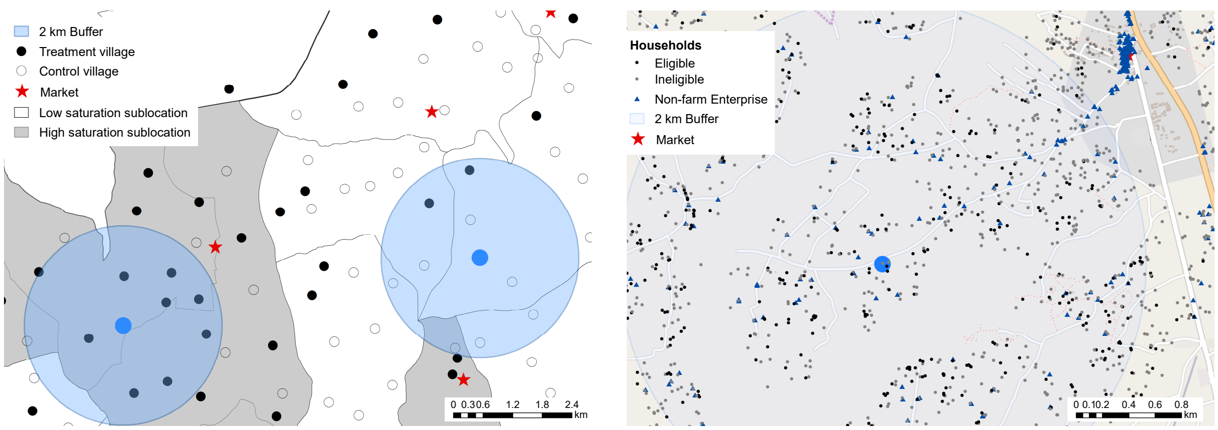
We then randomly assigned villages to three groups, and randomly assigned these groups to either a) treatment, b) treatment in high saturation, control in low saturation, and c) control. In Alego, 12 saturation groups were assigned to high saturation status, covering 98 villages (65 treatment, 33 control), and 11 saturation groups were assigned to low saturation status, covering 105 villages (37 treatment, 68 control). Across these 203 villages, a total of 7,891 households were classified as eligible by the GE census team (37 percent of households), with 3,950 of these households in treatment villages. We randomly generated an order for GD program expansion by first randomly ordering the saturation groups, and then villages within saturation groups.

The second randomization batch included villages in Ugunja and Ugenya subcounties. GD had not previously worked in any villages in these subcounties, so we did not stratify on any variables related to previous exposure for these villages. These subcounties had a larger number of villages per sublocation than Alego on average; as a result, we randomized saturation at the sublocation level. We assigned villages to one of three groups as above, pooled the "residual" villages that were not a multiple of 3, and randomly assigned one third of these to treatment, one third to treatment in high saturation sublocations and control in low saturation sublocations, and one third to control. GD worked first in Ugunja and then Ugenya. Across Ugunja and Ugenya, 115 sublocations covering 227 villages (148 treatment, 79 control) were assigned to high saturation status, while 79 sublocations covering 224 villages (78 treatment, 146 control) were assigned to low saturation status. These 450 villages had a total of 13,846 households classified as eligible by the GE census team (31 percent), with 7,105 of these households in treatment villages. We generated a random order within these subcounties by first ordering locations (the administrative unit above the sublocation), then sublocations within the location, then villages within the sublocation.

B.3 Illustrating spatial variation in treatment

Figure B.1: Spatial variation of data and treatment

(a) Comparing 2 villages (b) Households and enterprises



Notes: This figure provides an example of the spatial variation that we use to identify spillover effects. Both panels provide zoomed-in views on a selection of villages from Figure A.2. Panel A illustrates variation in the density of treatment villages around 2 treated villages. It plots village centers for treatment (filled circles) and control (open circles) villages, as well as a 2 km radius around the village center. While both villages themselves are not treated, the share of treated villages around them varies considerably. Panel B zooms in on one of these villages and plots eligible and ineligible households, as well as non-farm enterprises. Market centers are plotted as red stars.

We construct the amount per capita GDP in each buffer around a village or market i ($Amt_{it,r}$) as the per capita transfers in each buffer r, divided by per capita GDP. We obtain amount transferred into each buffer r at time t from the GPS location of recipients, as well as information from GiveDirectly (GD) on transfers sent to each recipient. Per capita GDP is calculated as the population-weighted average, across all households in the study area, of our expenditure-based measure of GDP (see Section 5). To convert stock values into annual flows, we assume a 10% depreciation rate. This yields an average per capita GDP of 637 USD PPP (or 2727 USD PPP per household). For $Amt_{vt,r}^{\neg v}$, we exclude households in buffer r but located in the same village v.

The population in each radius band around each market or village is determined using the GPS location of each household in our baseline household census data. Each household is then multiplied by the average number of people per household from the baseline household survey. This provides a population measure for each village in our study sample. To account for villages not included in our sample, but within radii bands of study markets or villages, we take two approaches. First, in villages that were not part of our sample but where GD had worked previously, we use household GPS locations provided to us by GD. For areas which were neither in our sample nor had been visited by GD previously, we calculate the population by uniformly distributing the sublocation population from the 2009 Kenyan census, net of the population in study area or GD census villages, over the area of the sublocation that was not already covered by a village in our study or a village where GD had worked previously. Village areas are defined as convex hulls around GPS coordinates of all village households. 2009 Kenyan census numbers are inflated by the overall average population growth rate in Kenya between 2009 and 2014.

C Household data appendix

C.1 Construction of index outcomes

Our index variables are constructed from the following components:

1. Psychological well-being index: Weighted, standardized average of depression (10 question CES-D scale), happiness, life satisfaction, and perceived stress (PSS-4), appropriately signed so that positive values represent better psychological well-being.

2. Health index: weighted, standardized average of self-reported health (on a scale of 1 to 5), an index of indicators for common health indicators, and an indicator for whether the respondent has experienced a major health problem since the date of baseline surveys, appropriately signed so that positive values represent better health.

3. Food security index: weighted, standardized index of the number of days a) adults and b) children i) skipped or cut meals, ii) went to bed hungry, iii) went entire days without food out of the last 7 days, appropriately signed so that higher values represent better food security. The Children food security index is made up of the child-related food security questions.

4. Education index: weighted, standardized average of total education expenditure and proportion of school-aged children in school, appropriately signed so that higher values represent better education outcomes.

5. Female empowerment index: weighted, standardized average of a violence index and attitudes index, appropriately signed so that positive values reflect more female empowerment/less domestic violence. The violence index is calculated as from the frequency of physical, emotional as sexual violence over the last 6 months. The attitudes index is calculated from an index of male-oriented attitudes and an index on the justifiability of domestic violence.

6. Security index: a weighted, standardized index of the number of times victimized by i) theft or ii) assault, arson or witchcraft in the last 12 months, an indicator for experiencing but not reporting a crime, and an indicator for reporting to be worried about crime or safety in the neighborhood.

C.2 Tracking and attrition

We achieved high tracking rates at endline, reaching over 90 percent of both treatment and control households. To assess levels of attrition, and whether attrition at endline is affected by treatment status and hence might confound our results, we estimate Equation (1) using as an outcome an indicator r_{hvs} for whether household h in village v in sublocation s is observed at endline, and do this separately for eligible and ineligible households. We investigate whether this indicator of non-attrition varies with treatment status in Table C.1.

We observe high tracking rates of 90.3 and 90.8 in the two types of households, respectively, in low-saturation control villages. These rates are very similar in other villages and sublocations: We observe broadly insignificant treatment coefficients in both tables, suggesting that attrition does not systematically vary with treatment status. This result is robust to defining r_{hvs} as an indicator for being reached at both baseline and endline (Column 2). It is also robust to restricting the sample to only households reached at endline (Panel B) or only households surveyed at baseline (Panel C). The one significant coefficient is for ineligible households in high-saturation sublocations: these are significantly less likely to be reached twice (Panel A, Column 4).

Table C.1: Household survey tracking and attrition

	(1)	(2)	(3)	(4)
	Eligible		**Ineligible**	
	Surveyed at endline	Surveyed at baseline and endline	Surveyed at endline	Surveyed at baseline and endline
Panel A: All households targeted at endline				
Treatment Village	0.006	0.006	0.011	0.017
	(0.009)	(0.013)	(0.011)	(0.016)
High Saturation Sublocation	−0.003	−0.019	−0.015	−0.035*
	(0.009)	(0.015)	(0.011)	(0.018)
Control, Low Sat Mean (SD)	0.903	0.815	0.905	0.811
	(0.296)	(0.389)	(0.293)	(0.392)
Observations	5,992	5,992	3,121	3,121
Panel B: Among households surveyed at endline				
Treatment Village		0.001		0.008
		(0.011)		(0.015)
High Saturation Sublocation		−0.019		−0.024
		(0.014)		(0.017)
Control, Low Sat Mean (SD)		0.902		0.895
		(0.297)		(0.306)
Observations		5,385		2,822
Panel C: Among households surveyed at baseline				
Treatment Village	−0.004	−0.004	0.012	0.012
	(0.009)	(0.009)	(0.011)	(0.011)
High Saturation Sublocation	0.000	0.000	−0.021*	−0.021*
	(0.009)	(0.009)	(0.011)	(0.011)
Control, Low Sat Mean (SD)	0.927	0.927	0.933	0.933
	(0.260)	(0.260)	(0.251)	(0.251)
Observations	5,150	5,150	2,661	2,661

Notes: This table reports tracking and attrition rates for households, by classification as eligible or ineligible to receive GD transfers by GE project field staff. Each Column represents a regression of an indicator for being surveyed at endline, or at both baseline and endline on an indicator for being in a treatment village, and and an indicator for the saturation status of the sublocation. Panel A includes all eligible households that were targeted for endline surveys. Panel B looks at eligible households that completed endline surveys, and serves as our main analysis sample. Panel C looks at eligible households that completed baseline surveys, and provides information on households that attrited from baseline to endline. Standard errors are clustered at the village level. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

C.3 Baseline balance

We re-estimate our main specifications using baseline demographic and outcome data, following the format of Table 1.[51] Panel A presents household demographics, while Panel B presents outcomes from Tables 1 and 2 for which we have baseline data (we did not collect consumption expenditure data at baseline). We are generally balanced across a wide range of variables. In our main specifications, we include baseline values of the outcome variable as a control when available to improve statistical precision.

C.4 Household weights

We weight household-level analyses with inverse sampling probability weights to ensure results are representative of the full population. In each village, we have baseline census data that provides the total number of households, classified by transfer eligibility status (based on research team reports). We targeted 8 eligible households and 4 ineligible households for surveys at baseline, and at endline targeted households surveyed at baseline, as well as those targeted and missed at baseline. The number of eligible households varies across villages; we thus weight households surveyed at endline by the inverse of the share of eligible households surveyed within the village. We do the same for ineligible households.

For hourly earnings, land prices and household interest rates, we interact these household level weights with the number of hours worked, acres of land owned, and total loan amounts, respectively, to make price effects interpretable as unit price effects.

51. We pre-specified a different set of balance checks that did not incorporate spatial variation; these are available in Egger et al. (2020). These checks also show the experiment is well-balanced.

Table C.2: Household balance

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation mean (SD)
Panel A: Respondent demographics				
Female	0.02	0.02	−0.01	0.75
	(0.02)	(0.03)	(0.02)	(0.43)
Respondent aged 25 or older	0.00	0.01	−0.01	0.92
	(0.01)	(0.02)	(0.01)	(0.28)
Is married	0.03	0.02	0.01	0.50
	(0.02)	(0.03)	(0.04)	(0.50)
Completed primary school	0.02	0.02	0.05*	0.33
	(0.02)	(0.03)	(0.03)	(0.47)
Has child	0.01	0.02	0.04*	0.73
	(0.01)	(0.02)	(0.02)	(0.44)
Self-employed	−0.01	−0.01	0.00	0.28
	(0.02)	(0.02)	(0.03)	(0.45)
Employed in wage work	−0.02	−0.01	0.01	0.25
	(0.02)	(0.05)	(0.03)	(0.43)
Panel B: Household assets				
Assets (non-land, non-house), net borrowi	3.95	−16.49	−41.27	1,017.56
	(23.06)	(36.24)	(100.34)	(1,391.45)
Housing value	14.47	−22.64	−0.21	1,579.57
	(14.43)	(19.55)	(342.18)	(4,219.38)
Land value	−221.59	−237.81	−199.16	4,366.48
	(162.63)	(339.89)	(451.87)	(5,819.84)
Panel C: Household cash flow				
Household non-ag income, annualized	−4.60	29.66	−14.39	197.28
	(15.82)	(32.74)	(29.24)	(461.78)
Self-employment profits, annualized	2.26	10.08	−2.57	89.07
	(7.48)	(13.67)	(18.83)	(288.60)
Wage earnings, annualized	−10.27	8.47	−5.97	97.14
	(12.68)	(26.09)	(13.80)	(309.63)
Tax paid, annualized	1.97	3.40**	3.27	16.33
	(1.20)	(1.71)	(2.44)	(44.73)
Panel C: Input Prices				
Land price per acre	−52.49	199.16	268.08	3,303.33
	(94.75)	(168.56)	(262.23)	(2,985.13)
Acres of land owned	35.61	71.96	−0.32**	1.36
	(35.66)	(72.82)	(0.15)	(2.39)
Total loan amount	1.65	6.44	−3.90	54.14
	(3.17)	(4.79)	(12.37)	(162.42)

Notes: This table presents regression specifications from Table 1 using baseline demographic and outcome variables. We did not collect consumption expenditure data at baseline. We have 4,674 to 4,768 observations for columns 1 and 2 (3,962 for land price) and 4,696 to 4,831 observations for column 3 (4,201 for land price). * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

C.5 Constructing average effects from coefficient estimates

Table C.3 presents the coefficient estimates underlying our reported average effects shown in Table 1. These average effects are constructed using the average values of each of the regressors included in the selected specification, denoted $\bar{X}$ and presented in the bottom row. For instance, the recipient household total effect for household expenditure (Table 1, row 1, column 2) is found by multiplying the coefficient on the amount going into households' own village Amt_v (row 1, column 2) by the mean amount (relative to village GDP) going into own village $\overline{Amt_v}|i$ is an eligible recipient (last row, column 2) and adding the coefficient on the amount going to other villages within 0-2km $Amt_{v,2}^{\neg v}$ (row 1, column 3) times the mean amount going into this radii band for treated villages $\overline{Amt}_{v,2}^{\neg v}$ (last row, column 3). We proceed in the same manner for the other tables.

Table C.3: Coefficient estimates for Expenditures, Savings and Income

	Recipient households			Non-recipient households		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$(Treat village) Reduced form	Amt Own Village IV	Amt Other Villages 0-2km IV	Amount, Control Eligibles 0-2km IV	Amount, Ineligibles 0-2km IV	Control, low-saturation mean (SD)
Panel A: Expenditure						
Household expenditure, annualized	292.98***	1,159.08***	375.10	265.50	4,419.08***	2,536.86
	(60.09)	(230.42)	(1,024.59)	(1,053.14)	(1,592.72)	(1,934.09)
Non-durable expenditure, annualized	186.96***	753.99***	318.11	311.50	4,180.12***	2,471.49
	(58.55)	(213.51)	(936.58)	(993.89)	(1,560.89)	(1,877.82)
Food expenditure, annualized	71.61*	352.01***	451.97	133.84	1,753.13**	1,578.43
	(36.93)	(128.50)	(584.12)	(629.77)	(767.95)	(1,072.31)
Temptation goods expenditure, annualized	6.51	23.71	−3.65	133.92	−37.57	37.10
	(5.79)	(22.07)	(94.11)	(100.81)	(83.98)	(123.59)
Durable expenditure, annualized	95.18***	368.21***	137.16	71.42	97.73	59.44
	(12.64)	(57.53)	(194.97)	(211.63)	(161.52)	(230.90)
Panel B: Assets						
Assets (non-land, non-house), net borrowing	178.09***	674.84***	55.46	−155.73	1,811.04*	1,132.15
	(24.61)	(87.84)	(376.80)	(501.75)	(1,055.96)	(1,420.22)
Housing value	372.78***	1,456.55***	1,082.00***	341.81	901.54	2,033.72
	(25.25)	(87.55)	(354.50)	(469.08)	(2,888.39)	(5,030.37)
Land value	50.86	330.17	728.74	2,340.15	7,179.88	5,030.72
	(186.08)	(478.95)	(2,920.78)	(3,650.75)	(5,854.60)	(6,607.61)
Panel C: Household balance sheet						
Household income, annualized	77.62*	333.58*	510.04	1,021.65	2,862.47**	1,023.45
	(43.66)	(185.90)	(746.45)	(746.49)	(1,164.68)	(1,634.70)
Net value of household transfers received, annualized	−1.68	−5.94	−64.41	−85.85	135.35	130.18
	(6.81)	(25.92)	(115.95)	(129.15)	(249.65)	(263.75)
Tax paid, annualized	1.95	7.31	−21.99	−11.57	24.72	16.93
	(1.28)	(4.50)	(19.59)	(20.76)	(25.79)	(36.51)
Profits (ag & non-ag), annualized	24.70	115.74	37.09	−78.33	607.47	485.20
	(23.18)	(97.59)	(388.53)	(477.68)	(606.86)	(787.10)
Wage earnings, annualized	42.51	196.44	243.31	1,131.87**	2,216.33**	495.37
	(32.24)	(119.53)	(557.31)	(491.88)	(862.71)	(1,231.56)
$\bar{X}$		0.26	0.09	0.08	0.09	

Notes: This table reports the coefficient estimates that underlie the average effects reported in Table 1, see corresponding table note for more details. $\bar{Amt}$ reports the average of each RHS variable for the sample studied (recipients or non-recipients), which we multiply with the coefficient to get the average effects reported.

D Enterprise data appendix

D.1 Enterprise census and survey details

We conducted a baseline enterprise census in each village on the same day as the baseline household census. The household census included a question on whether the household was running an enterprise from their homestead or from a fixed kiosk/shop. The enterprise census targeted enterprises operating outside of homesteads. We then returned to survey enterprises operating outside of the homestead and open on the day of our visit, coincident with baseline household surveys. In villages with over 20 enterprises operating outside of homesteads, e.g., those that overlapped a market center, we randomly selected 20 enterprises to survey.

Our endline enterprise census sought to re-identify all enterprises operating from within or outside homesteads, both those identified at baseline and any new enterprises. In order to maintain a representative sample, we randomly sampled up to 2 enterprises operating from within homesteads and up to 3 outside of homesteads to be surveyed, including those in market centers in villages containing a market.

Enterprise surveys cover profits, revenues, and a subset of costs (including the wage bill), and at endline collected information on inventories and investment. We measure (annualized) revenues and profits for non-agricultural enterprises directly by asking respondents about these quantities with a one month recall period (de Mel, McKenzie, and Woodruff 2009). We calculate costs as the sum of the employee wage bill, rent and security costs; this is not a comprehensive measure of all costs, and hence we do not expect the revenue measure to equal our measure of profits plus measured costs. In particular, we do not directly measure expenditure on intermediate inputs such as materials or supplies.

Information on agricultural enterprises comes from our household surveys. Baseline household surveys did not include sufficient detail to construct measures of agricultural revenue and profit, so we only use endline measures for these outcomes. For agricultural enterprises, total revenue is calculated as the sum of crop output (measured at the crop level) plus the value of pastoral and poultry output sold, and the value of the household's own consumption of pastoral and poultry output. When crop output was reported in non-monetary units, we convert these to monetary values using the 2016 mean of the median crop output price measured in the market price surveys in the household's sub-county. Agricultural costs are the wage bill, all agricultural inputs (e.g., seed and fertilizer), and land rental costs. We then calculate agricultural profits as total agricultural revenue minus agricultural costs.

D.2 Enterprise specifications

We estimate the following equations for enterprises:

$$y_{ivs} = \alpha_1 Treat_v \cdot X_{ivs} + \alpha_2 HighSat_s \cdot X_{ivs} + X_{ivs}\gamma + \delta_1 y_{ivs,t=0} \cdot X_{ivs} + \delta_2 M_{ivs} \cdot X_{ivs} + \varepsilon_{ivs}, \tag{10}$$

$$y_{iv} = \beta Amt_v \cdot X_{ivs} + \sum_{r=2}^{R} \beta_r Amt_{v,r}^{\neg v} \cdot X_{ivs} + X_{ivs}\gamma + \delta_1 \bar{y}_{iv,t=0} \cdot X_{ivs} + \delta_2 M_{iv} \cdot X_{ivs} + \varepsilon_{iv}. \tag{11}$$

Here, y_{ivs} is an outcome for enterprise i in village v (and sublocation s), $X_{iv(s)}$ is a vector of indicators for enterprise type (agricultural, non-agricultural operating outside the homestead, non-agricultural operating from the homestead), and other terms are defined as in Section 3. We interact our treatment indicator and transfer amount variables with this vector of enterprise types, effectively estimating a stacked version of Equations 1 and 2. This allows treatment effects and controls to vary flexibly across enterprise type. Table D.1 reports the share of enterprises by sector weighted by count and by revenue. Since enterprise surveys were conducted as repeated cross-section rather than a panel, we control for the village-level baseline mean of the outcome variable where available in our main specification. Results are similar if we omit this control (Table D.2).

We use our endline enterprise census data to construct weights that are representative of the full population of enterprises. In particular, we weight enterprises by the inverse of the share of surveyed enterprises of a particular type (agricultural, operating from homesteads, operating outside homesteads) within each village. For hourly wages, we interact these enterprise-level weights with the total hours worked to make wage effects interpretable as the average effect per hour worked.

D.3 Tracking, balance and attrition

Our enterprise samples are repeated cross-sections, so we do not report attrition rates between baseline and endline. We do check baseline balance for enterprises, taking the same approach as in Table 3 but using baseline values for outcomes that are available. (We did not collect enterprise investment or inventories, nor do we have revenue and profit measures for non-agricultural enterprises at baseline.) The baseline sample generally appears balanced; there are no statistically significant differences at the 5% level (Table D.3).

D.4 Matching enterprise owners

Through our integrated approach to enterprise and household censusing, we are able to match all agricultural enterprises (as found via household surveys), and 56% of non-agricultural enterprises, for a total of 93% of all enterprises. To match non-agricultural enterprises to the households that own them we apply both automatic and manual procedures to our detailed name, phone number and GPS data. As we relied heavily on the reported operating location, we excluded enterprise census data without this information. The proportion of matched enterprises are relatively evenly split by treatment status for both eligible and ineligible households: 52% of matched eligible enterprise owners and 51% of matched ineligible owners are in treatment villages.

Patterns with respect to the eligibility status of the owner are generally sensible: 28% of non-agricultural enterprises are owned by an eligible household, slightly below their share in the population (33%), and enterprises owned by ineligible (and thus on average somewhat richer) households have 9% higher profits and 21% higher revenues on average than those owned by eligibles.

Table D.1: Composition of enterprises by sector

	Overall		**Non-Ag**	
Sector	Count Share	Revenue Share	Count Share	Revenue Share
Retail	0.09	0.34	0.54	0.52
Manufacturing	0.04	0.16	0.24	0.24
Services	0.03	0.16	0.21	0.24
Agriculture	0.84	0.34		

Notes: This table describes enterprise shares by sector, both in terms of counts and shares of total revenue. Data on counts comes from the endline enterprise census (for non-agricultural enterprises) and the baseline household census (for agricultural enterprises). Data on revenue shares for the non-agricultural sectors comes from endline enterprise surveys, while data on agricultural revenue shares comes from endline household surveys.

Table D.2: Enterprise outcomes without baseline controls

	(1)	(2)	(3)	(4)
	Treatment Villages		**Control Villages**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Panel A: All enterprises				
Enterprise profits, annualized	0.02	63.01	36.52	323.39
	(30.07)	(163.80)	(141.74)	(691.12)
Enterprise revenue, annualized	−82.06	376.54**	303.94**	758.52
	(129.11)	(148.36)	(139.85)	(2,493.40)
Enterprise costs, annualized	−12.61	106.19**	78.85	147.73
	(34.91)	(42.42)	(53.53)	(550.11)
Enterprise wagebill, annualized	−15.49	91.51***	77.86*	120.62
	(30.84)	(33.73)	(40.73)	(492.11)
Enterprise profit margin	—	—	—	—
Panel B: Non-agricultural enterprises				
Enterprise inventory	—	—	—	—
Enterprise investment, annualized	—	—	—	—
Panel C: Village-level				
Number of enterprises	0.01	0.01	0.00	1.12
	(0.01)	(0.02)	(0.02)	(0.14)

Notes: This table replicates Table 3 but without village level baseline control variables. We omit outcomes for which baseline controls were not available in the original table, as results for those outcomes are unaffected. See notes to Table 3 for further details. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table D.3: Enterprise Balance

	(1)	(2)	(3)	(4)
	Treatment Villages		**Control Villages**	
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Control, low saturation weighted mean (SD)
Panel A: Non-agricultural enterprises				
Enterprise profits, annualize	–10.29	6.96	12.99	1,141.19
	(16.29)	(21.84)	(20.63)	(1,848.50)
Enterprise revenue, annualized	–99.79	65.88	118.63	4,919.61
	(87.63)	(99.52)	(103.40)	(11,430.88)
Panel B: All enterprises				
Enterprise costs, annualized	3.78	16.15	6.79	57.09
	(6.05)	(10.37)	(9.32)	(237.07)
Enterprise wagebill, annualized	3.39	15.75	6.95	53.08
	(5.91)	(10.00)	(8.94)	(233.12)
Panel C: Village-level				
Number of enterprises	0.00	–0.01	–0.01	1.07
	(0.01)	(0.02)	(0.02)	(0.14)

Notes: This table presents regression specifications from Table 3 using corresponding baseline enterprise outcomes where available. We did not collect enterprise inventories and investment data at baseline. We also exclude baseline agricultural revenues and profits, as these were not collected in the same manner as at endline. We have between 4,125 and 4,193 observations in Panel A, 9,245 to 9,264 in Panel B, and 653 in Panel C. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct.

E Price data appendix

E.1 Categorizing market survey products

Our market surveys included questions about 84 commodities. As outlined in our pre-analysis plan, we excluded products that are not present in at least 5 percent of market-month observations; this affects 11 products (bicycle, bull (grade), calf (grade), donkey, duck, piglet, turkey, goat (meat), lamb, milk powder, and mosquito net). Three products (Waterguard, fertilizer, and improved maize seeds) do not have consumption expenditure analogues in the KLPS-3 surveys that we use to construct expenditure weights, so these are also excluded, leaving a final list of 70 products.

Table E.1 presents the classification of the products we use in our price analysis into more and less tradable categories, and the subcategories (food, livestock, (non-food) non-durables, durables and temptation goods) shown in Table 4.

Table E.1: List of market products by category

Less tradable (locally produced)			**More tradable**		
Food		Livestock	Non-Food Non-Durables	Durables	Temptation Goods
Cassava	Papaya	Bull (local)	Bar soap	1 Iron sheet (32 gauge)	Cigarettes
Irish potato	Pineapple	Calf (local)	Toothpaste	Cement	
Maize	Water Melon	Chicken (hen)	Vaseline/lotion	Large Padlock	
Millet	Jackfruit	Goat	Washing powder	Nails (3 inch)	
Plantains	Passion Fruit	Sheep	Bleach	Roofing Nails	
Rice	Beef		Panadol/aspirin	Timber (2x2)	
Sorghum	Fish (Tilapia)		Cooking fat	Water Paint	
Sweet potato	Pork		Batteries (3-volt)	20L Jerry can	
Beans	Eggs		Firewood	Thermos flask	
Cabbage	Milk (Fresh)		Kerosene	3 1/2 X 6 Mattress	
Cowpea leaves	Biscuits		Charcoal		
Green grams	Bread		Leso		
Groundnuts	Cake		Small sufuria		
Kales	Maize flour		Slippers		
Onions	Wheat flour				
Saka (Local Vegetable)	Milk (Fermented)				
Tomatoes	Soda				
Avocado	Sugar				
Banana-sweet	Tea				
Mango					
Orange					

Notes: This table presents the classification of the 70 products used in our analysis of output prices. The classification follows our midline pre-analysis plan (Appendix I). The market survey collected information on 85 products. As outlined in our pre-analysis plan, we exclude any product that, at the market-product level, is missing for more than 95% of cases, a total of 11 products. We also drop three products that do not match items in our expenditure share data.

E.2 Price analyses robustness checks

E.2.1 Alternative definition of market access

Our main specifications separates price effects by market access as defined in Donaldson and Hornbeck 2016: $MA_m = \sum_{r=0}^{10} r^{-\theta} N_r$, where $\theta = 8$ and N_r is the population in in the $r - 2$ to r km buffer around each market. Here we present alternative results based on a definition of market access as the inverse distance from the closest 'main' road, where we define a main road as any road in Open Street Maps classified as motorway, trunk, primary, secondary or tertiary road (excluding residential streets, tracks, paths, and unclassified roads). While price effects were concentrated in low-market-access areas using our main population-density-based market access measure, they seem to be fairly similarly small when splitting by road access. In a context where most people walk to their nearest market, this may not be surprising. However, we cannot reject that results are the same as our main results.

Table E.2: Output Prices using distance to main road as market access measure

		(1)	(2)	(3)	(4)
		Overall Effects		**ATE by road access**	
		ATE	Average maximum effect (AME)	below median	above median
All goods		0.0010*	0.0042	0.0010	0.0011
		(0.0006)	(0.0031)	(0.0008)	(0.0008)
By tradability	More tradable	0.0014	0.0062	0.0006	0.0021
		(0.0015)	(0.0082)	(0.0021)	(0.0021)
	Less tradable	0.0009	0.0034	0.0012	0.0007
		(0.0006)	(0.0032)	(0.0009)	(0.0009)
By sector	Food items	0.0009	0.0036	0.0014	0.0007
		(0.0006)	(0.0033)	(0.0009)	(0.0010)
	Non-durables	0.0014	0.0061	0.0005	0.0020
		(0.0017)	(0.0089)	(0.0023)	(0.0022)
	Durables	0.0019*	0.0070	0.0012	0.0031
		(0.0011)	(0.0061)	(0.0013)	(0.0019)
	Livestock	−0.0008	−0.0027	−0.0023*	0.0012
		(0.0010)	(0.0052)	(0.0013)	(0.0013)
	Temptation goods	−0.0011	−0.0112	−0.0035	0.0022
		(0.0026)	(0.0143)	(0.0036)	(0.0041)

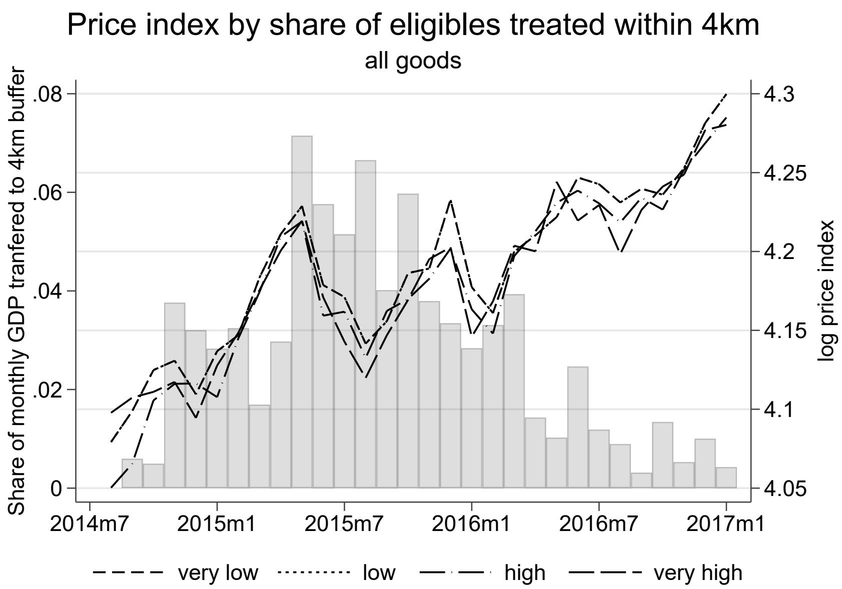
Notes: This table replicates Table 4. See notes for details. The only difference is the definition of market access of each market in Columns 3 and 4: It is defined as the inverse distance from the closest main road, classified by Open Street Map as motorway, trunk, primary, secondary or tertiary road (excluding residential, tracks, paths, and unclassified roads): $MA_m = \frac{1}{distance_m}$. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

E.2.2 Spatial and temporal analysis of price effects

Our main analysis in Section 4.5 follows our pre-specified algorithm, which selects the number of lags and distance buffers by minimizing the Schwarz BIC. While we allow for up to 18 months lags, and 20km spatial dependence, the algorithm selects a specification that includes only contemporaneous transfers up to 4km. In this section, we present three pieces of additional exploratory analysis that serve as robustness checks on our primary pre-specified results and explore the spatial and temporal dimensions of price effects in turn.

First, we show our price data in raw form: Figure E.1 shows that prices in more vs. less exposed markets as measured by the share of eligible households within 4km that were assigned to treatment evolved very similarly over the course of the study period and afterwards. We can visually reject large differences in the evolution of prices in response to treatment.

Figure E.1: Price index by treatment intensity



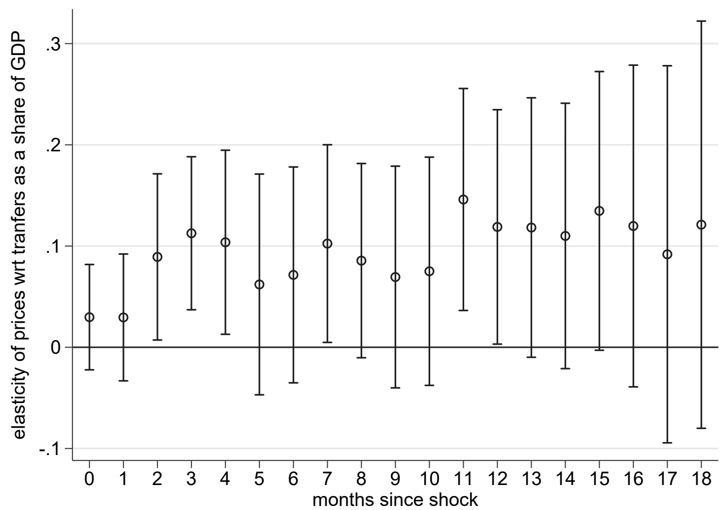
Notes: The figure shows the log price index across all goods in more vs. less exposed markets as measured by quartiles of the the share of eligible households within 4km that were assigned to treatment. Bars represent average transfer amounts relative to monthly GDP going to the 4km buffer across all 61 markets in each month.

Second, we estimate Equation (4) for a range of outer radii R from 2km to 6km while fixing the number of temporal lags at the (BIC selected) value of 0 months. This allows us to test whether our algorithm is indeed picking up the relevant spatial horizon, or whether we might be missing part of the effect. Table E.3 shows that price effects are robust to including additional radii bands. For none of the price indices can we reject that adding an additional buffer on top of that selected by our pre-specified algorithm leads to significantly different average price effects as those in the main specification.

Third, we do the analogous exercise temporally, estimating Equation (4) for lag structures of up to $M = 18$ months while fixing the maximum radius at $R = 4$km. We then calculate the cumulative effect of a shock of 100% of monthly GDP in each month up to L months. Figure E.2 shows that prices adjust rapidly. We cannot reject that lags of treatment beyond 3 months have no additional effect on prices, and the elasticity of prices with respect to transfers as a share of GDP stabilizes around 0.1.

The economic implication is that any part of the price response we miss using our pre-specified algorithm to select a temporal horizon does not make a meaningful difference quantitatively. In the most intense 12 month period, 3.9% of annual GDP was transferred to the 0 - 4km buffer around the average market (see Figure E.1). With an elasticity of 0.1, this implies a price effect of 0.4% in the average market over the most intense period of transfers (nearly identical to the 0.4% we arrive at using our pre-specified algorithm).

Figure E.2: Cumulative price effects



Notes: The figure is based on estimating Equation 5, where we impose a maximum lag or price effects up to $M = 18$ months and a maximum spatial radius of $R = 4$km. We then calculate the cumulative effect of a shock of 100% of monthly GDP in each month over L months on the overall logarithmic price index ($= \sum_{r=2km}^{R=4km} \sum_{l=0}^{L} \hat{\beta}_{rl}.$). Confidence intervals are as in Conley (2008) and we allow for spatial correlation up to 10km and autocorrelation up to 12 months.

E.2.3 IV specification for market price effects

In our preferred estimates we identify effects on market prices using a different method than the IV strategy we use for identifying effects on households and firms (i.e. specifications 2, 3 and 11). This is because, unlike for households and firms, we have rich panel data on market prices including prices both before and after treatment onset. This lets us identify effects using a difference-in-differences design (as described in section 3.7) and leveraging the random roll-out of transfers into different buffers around each market over time. Specifically, conditional on market fixed effects, which control for the share of eligible households around each market as well as other time-invariant characteristics of markets, treatment roll-out is exogenous.

For comparability, however, we can also estimate price effects using the same IV strategy we use for firm and household outcomes. Concretely, we run the same pre-specified radii and buffer selection algorithm (specifications 4 and 5) based on the BIC for each price effect. But instead of including market fixed effects and $Amt_{m(t-l),r}$, we drop market fixed effects, and instrument $Amt_{m(t-l),r}$ with share $s_{m,r}^{e,t}$ of eligible households assigned to treatment in the buffer r around market m multiplied by the share of transfers going to that buffer in month $t-l$. Note that this approach is analogous to our dynamic IV specification 7 we use to estimate the impulse response functions for flow variables underlying our multiplier estimates. After selecting radii $\bar{R}$ and lags L to be included for each price index, we run:

$$p_{mt} = \sum_{r=2}^{\bar{R}} \sum_{l=0}^{L} \beta_{rl} Amt_{m(t-l),r} + \lambda_t + \varepsilon_{mt} \tag{12}$$

Table E.4 reports average and average maximum effects as in Table 4 resulting from this strategy. Effects are broadly in line with those in our main specification. Although none of the effects are statistically significant when using the IV strategy, we are still able to reject large effects on prices.

E.3 Enterprise price analyses

In addition to prices collected as part of our market price surveys, we also collected some price data as part of our enterprise surveys. We make use of enterprise price data collected via seven rounds of phone surveys of enterprises between August 2015 and June 2016. These surveys were conducted with four types of enterprises: small retailers, hardware stores, maize grinders, and tailors. We focus on prices for services provided by the latter two, as hardware and retail prices are well-covered by our market price data. To ensure consistent quality, unit size and availability we collected prices for a small number of services these enterprises commonly provide. In particular, we focus on the price of grinding 1kg of maize at a posho mill, and for patching a small hole at a tailor shop.

Phone surveys overlapped with an intense period of treatment rollout. During those 11 months the share of overall transfers sent went from 52% to 92%, and the variation in transfers was substantial, both across space and time: The 10-90 percentile range of per capita GDP transferred within 2km of a village over the period is $[0.5\%, 9.6\%]$, and the average village experienced 1.8% of GDP more inflows in the most intense month compared to the least intense month.

We analyze these prices analogously to our market prices, running the following specification:

$$p_{evt} = \sum_{r}^{\bar{R}} \sum_{l=0}^{M} \beta_{rl} Amt_{v(t-l),r} + \alpha_v + \lambda_t + \varepsilon_{evt} \tag{13}$$

where p_{evt} is the logarithm of the price from enterprise e in village v in month t, α_v are village fixed effects, γ_t are month fixed effects. We select the included radii bands $\bar{R}$ and the number of treatment lags M using the same pre-specified algorithm as for market prices. Table E.5 reports the average treatment effect across the intervention period (ATE) as well as the average maximum effect across villages (AME) from the optimal specification, and investigates heterogeneity by market access (see Section 3.7 for details on the methodology).

We find limited effects on these two selected services, with magnitudes in the range of product-specific effects for our our market price measures. Tailoring prices rise by 0.02% on average, and 0.1% in the month of most intense transfer, though those coefficients are not statistically significant. As with market prices, the effects are concentrated in more remote areas. Maize grinding prices fall, if anything, but the estimated effects are not statistically significantly different from zero.

Table E.3: Robustness to fixing alternative radii bands: Output Prices

		Overall Effects			
		(1) ATE Optimal Radius	(2) ATE $\bar{R} = 2$	(3) ATE $\bar{R} = 4$	(4) ATE $\bar{R} = 6$
All goods		0.0010*	0.0001	0.0010*	0.0014*
		(0.0006)	(0.0004)	(0.0006)	(0.0008)
By tradability	More tradable	0.0014	0.0003	0.0014	0.0021
		(0.0015)	(0.0010)	(0.0015)	(0.0020)
	Less tradable	0.0009	0.0001	0.0009	0.0012
		(0.0006)	(0.0004)	(0.0006)	(0.0008)
By sector	Food items	0.0009	0.0001	0.0009	0.0012
		(0.0006)	(0.0005)	(0.0006)	(0.0009)
	Non-durables	0.0014	0.0003	0.0014	0.0021
		(0.0017)	(0.0011)	(0.0017)	(0.0021)
	Durables	0.0019*	0.0000	0.0019*	0.0027
		(0.0011)	(0.0008)	(0.0011)	(0.0016)
	Livestock	−0.0008	0.0001	−0.0008	−0.0011
		(0.0010)	(0.0006)	(0.0010)	(0.0011)
	Temptation goods	−0.0011	−0.0022	−0.0011	0.0002
		(0.0026)	(0.0019)	(0.0026)	(0.0034)

Notes: This table replicates Column 1 of Table 4, and then estimates the same ATE based on specifications where the maximum radius is imposed to be at $R \in [2\text{km}, 6\text{km}]$. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table E.4: Output Prices - IV Specification

		(1)	(2)	(3)	(4)
		Overall Effects		**ATE by market access**	
		ATE	Average maximum effect (AME)	below median	above median
All goods		0.0005	0.0035	0.0007	0.0004
		(0.0007)	(0.0050)	(0.0008)	(0.0010)
By tradability	More tradable	0.0002	0.0018	0.0010	0.0003
		(0.0012)	(0.0088)	(0.0019)	(0.0014)
	Less tradable	0.0006	0.0043	0.0006	0.0004
		(0.0008)	(0.0057)	(0.0011)	(0.0012)
By sector	Food items	0.0006	0.0046	0.0007	0.0005
		(0.0008)	(0.0060)	(0.0011)	(0.0012)
	Non-durables	0.0003	0.0019	0.0013	0.0003
		(0.0013)	(0.0094)	(0.0020)	(0.0015)
	Durables	0.0001	0.0006	−0.0021	0.0013
		(0.0015)	(0.0112)	(0.0025)	(0.0021)
	Livestock	−0.0005	−0.0038	0.0001	−0.0015
		(0.0009)	(0.0067)	(0.0006)	(0.0017)
	Temptation goods	−0.0029	−0.0214	−0.0033	−0.0029
		(0.0020)	(0.0147)	(0.0028)	(0.0022)

Notes: Each row represents an IV regression of the logarithm of a price index on the “optimal” number of lags and distance buffers of per capita Give Directly transfers in each buffer (Equation 12). Price indices are based on 311,138 non-missing price quotes for 70 commodities and products. For each product, we take the logarithm of the median price quote in a market-month, and create our market price indices as an expenditure weighted average of these median price quotes across all goods in that market-month. Regressions include a panel of 1,734 market-by-month observations. The number of radii bands and lags is chosen sequentially by minimizing the BIC, as pre-specified, for the overall price index, which selects 2km. Treatment amounts for each buffer-month are instrumented by the share of eligible households assigned to treatment in that buffer, multiplied by the share of all transfers in that buffer going out in that month. Regressions include a full set of month fixed effects. Column 1 reports the implied ATE, calculated by evaluating the “optimal” regression specification at the average level of treatment intensity between September 2014 and March 2017, the time during which transfers went out. Column 2 reports the average maximum effect, calculated at the average across all markets of the month in which the largest per capita transfers went into a market’s neighborhood (up to the largest buffer selected by the algorithm). Columns 3 and 4 break down the ATE by market access, defined as $MA_m = \sum_{r=1}^{10} r^{-\theta} N_r$, where $\theta = 8$ and N_r is the population in in the $r-2$ to r km buffer around each market. Standard errors (in parentheses) are as in Conley (2008) and we allow for spatial correlation up to 10km and autocorrelation up to 12 months. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table E.5: Local manufacturing and services prices

	(1)	(2)	(3)	(4)
	Overall Effects		**ATE by market access**	
	ATE	Average maximum effect (AME)	below median	above median
Tailor, patch small hole	0.0002	0.0011	0.0042	−0.0020
	(0.0034)	(0.0192)	(0.0033)	(0.0046)
Posho mill, grind 1kg of maize	−0.0011	−0.0061	−0.0022	−0.0010
	(0.0017)	(0.0097)	(0.0026)	(0.0014)

Notes: Each row represents a regression of the logarithm of a price on the "optimal" number of lags and distance buffers of per capita GiveDirectly transfers in each buffer (Equation13). We include 2,347 monthly price observations for tailors (simple patch), and 4,577 observations from posho mills (grinding 1kg of maize) collected between Aug 15 - Jun 16, around the time of peak transfer intensity. The number of radii bands and lags is chosen sequentially by minimizing the BIC, as pre-specified. Regressions include a full set of market and month fixed effects. Column 1 reports the implied ATE, calculated by evaluating the "optimal" regression specification at the average level of treatment intensity between September 2014 and March 2017, the time during which transfers went out. Column 2 reports the average maximum effect, calculated at the average across all markets of the month in which the largest per capita transfers went into a market's neighborhood (up to the largest buffer selected by the algorithm). Columns 3 and 4 break down the ATE by market access, defined as $MA_m = \sum_{r=1}^{10} r^{-\theta} N_r$, where $\theta = 8$ and N_r is the population in in the $r - 2$ to r km buffer around each market. Regressions are weighted by inverse sampling weights. Standard errors (in parentheses) are as in Conley (2008) and we allow for spatial correlation up to 10km and autocorrelation up to 3 months. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

F Robustness to alternative spatial modelling approaches

In this section we examine the robustness of our statistical inferences and overall conclusions to several alternative ways of dealing with the issue of radius selection in our spatial models.

F.1 Fixed radii

We first examine results holding the spatial radius fixed at 2km (the maximal radius our BIC procedure usually selects) as well as at 4km and 6km. Generally speaking we expect to capture additional spillovers using the larger radii that we might miss at the shorter 2km radius, but at the cost of precision. Tables F.1, F.2, and F.3 mirror our main Tables 1, 2 and 3 but using this approach. In each table Column 1 reproduces our main estimate while Columns 2, 3 & 4 report estimates using fixed 2km, 4km and 6km radii, all for recipient households (or in Table F.3 for treated villages). Columns 5-8 then repeat this exercise for non-recipient households and control villages.

At a fixed 2km radius our results are (not surprisingly) similar if not identical to those from our default specification, but with the advantage that inference post-model selection is less of a concern. At higher radii the point estimates are generally similar to (or in some cases larger than) our benchmark estimates, though as expected the precision of our estimates decreases at higher radii. In almost all cases we cannot reject that the fixed-radius estimates are the same as our benchmark estimates.

F.2 Split-sample estimation

We next examine robustness to selecting a radius and estimating coefficients using different splits of the data. Specifically, we select 200 random 50-50 splits of our data, stratified by village treatment assignment and (for households) eligibility status, into training and estimation samples. For each split we use the training sample to select a radius and the estimation sample to estimate parameters. We repeat this exercise, using the estimation sample as the training sample and vice-versa. We record the proportion of splits in which we calculate the same optimal radius band as when using our full dataset; we take the mean of the two point estimates and report the proportion of cases in which the resulting estimate effect lies within the 95% confidence interval reported in the paper.

Tables F.4, F.5, and F.6 presents results for the outcomes found in Tables 1, 2 and 3, respectively. For Tables F.4 and F.5, columns 1 and 2 reproduce the estimates and radii selection for recipient households. Column 3 reports the fraction of these splits that produce estimates for non-recipients falling into the 95%-CI of the initial estimate, and Column 4 the proportion that select the same radius as when using the full dataset. Columns 5-8 do the same for non-recipient households. Note that as in producing our main estimates we do not separately estimate an optimal radius for subcomponents of larger totals or indices. For enterprises, TableF.6 columns 1 and 3 reproduce the main estimates for treatment and control villages, respective. Columns 2 and 4 report the share of mean estimates falling within the 95%-CI of the initial estimate. As we use a common radius for treatment and control villages, column 5 reports the selected radius, and column 6 reports the share selecting the same radius.

Overall we see congruence between the full data and the subsamples regarding the optimal radius over which to estimate effects, with most agreement rates in the 90%s. We also see good coverage, with 95% or more of the mean replicate point estimates falling within our original 95% confidence interval in most cases.

F.3 Heterogeneous radii

We next examine whether our BIC algorithm selects different maximal radii for different geographic sub-groups of villages. Specifically, we (i) allow the BIC to select a different radius for markets with above versus below median market access, and then (ii) allow the BIC to select a different radius for each of the three sub-counties in which our study is set. Tables F.7, F.8 and F.9 report the maximal radius selected in each case, with the full sample radius selected for comparison. For enterprise results, optimal radii bands were selected only once for each outcome across treatment and control villages, as the enterprises were not direct recipients of the cash transfers. In the great majority of cases we end up selecting the same radius (which in almost every case is 2km). Specifically, out of 190 radii selected (15 outcomes * 2 treatment status groups * 5 geographic subgroups for households + 8 outcomes * 5 geographic subgroups for enterprises) we select a different radius than in the corresponding pooled approach 10 times, or 5.3% of the total. Overall we conclude that, while there are surely are differences in the relevant radii or more generally the relevant "catchment areas" for different units, our data do not reveal systematic differences.

F.4 Randomization Inference

Finally, we examine the sensitivity of our conclusions to randomization inference. This approach sidesteps concerns about model selection; we simply interpret the coefficients we obtain from the entire model selection and estimation procedure as a statistic whose distribution should be invariant to reassignments of treatment and control status under the null of no treatment effects for any unit. Specifically, we generate 500 replicates in each of which we re-assign each village and household's treatment status using the same algorithm with which actual treatment was assigned, recalculate the our derived spatial exposure measures using these assignments, and then re-estimate total effects.

Tables F.10, F.11 and F.12 report results for the outcomes in Tables 1, 2 and 3, respectively. Table F.13 does the same but also simulating the randomized rollout of the transfer program in order to conduct randomization inference for output price outcomes in Table 4. Randomization inference yields very similar substantive conclusions to our main analysis, rejecting the null of no treatment effects for almost exactly the same outcomes as our main tests reject the null of no average effect.

Table F.1: Robustness to fixing alternative radii bands: Expenditures, Savings and Income

	Recipient households				Non-recipient households				
	(1) Total Effect IV Optimal Radius	(2) Total Effect IV $\bar{R} = 2$	(3) Total Effect IV $\bar{R} = 4$	(4) Total Effect IV $\bar{R} = 6$	(5) Total Effect IV Optimal Radius	(6) Total Effect IV $\bar{R} = 2$	(7) Total Effect IV $\bar{R} = 4$	(8) Total Effect IV $\bar{R} = 6$	(9) Control, low-saturation mean (SD)
Panel A: Expenditure									
Household expenditure, annualized	338.16***	338.16***	516.73***	644.26*	333.73***	333.73***	448.96**	−31.50	2,536.86
	(109.36)	(109.36)	(178.89)	(382.56)	(123.24)	(123.24)	(182.86)	(408.82)	(1,934.09)
Non-durable expenditure, annualized	226.74**	226.74**	459.92***	476.74	316.62***	316.62***	296.15**	235.08*	2,471.49
	(99.62)	(99.62)	(169.26)	(382.52)	(119.79)	(119.79)	(119.53)	(120.56)	(1,877.82)
Food expenditure, annualized	133.55**	133.55**	323.17***	220.53	132.84**	132.84**	123.56**	101.04*	1,578.43
	(63.98)	(63.98)	(94.08)	(197.79)	(58.58)	(58.58)	(59.18)	(60.55)	(1,072.31)
Temptation goods expenditure, annualized	5.88	5.88	14.88	28.63	−0.71	−0.71	−2.20	−5.09	37.10
	(8.82)	(8.82)	(11.87)	(25.23)	(6.50)	(6.50)	(6.54)	(7.27)	(123.59)
Durable expenditure, annualized	109.07***	109.07***	59.95***	147.41***	8.41	8.41	8.21	13.19	59.44
	(20.23)	(20.23)	(22.77)	(40.73)	(12.50)	(12.50)	(12.89)	(11.66)	(230.90)
Panel B: Assets									
Assets (non-land, non-house), net borrowing	182.01***	182.01***	204.08***	291.61**	132.63*	132.63*	244.56*	123.84	1,132.15
	(44.25)	(44.25)	(66.29)	(141.20)	(78.32)	(78.32)	(142.59)	(244.33)	(1,420.22)
Housing value	480.68***	480.68***	479.76***	595.96***	72.58	72.58	543.67	−402.46	2,033.72
	(38.88)	(38.88)	(48.66)	(103.95)	(215.70)	(215.70)	(474.16)	(845.95)	(5,030.37)
Land value	153.09	153.09	511.79	239.44	572.07	572.07	816.20	−261.31	5,030.72
	(262.48)	(262.48)	(357.15)	(672.45)	(458.28)	(458.28)	(728.50)	(1,477.50)	(6,607.61)
Panel C: Household balance sheet									
Household income, annualized	134.02	134.02	228.45	405.02	229.46***	229.46***	315.10**	171.59	1,023.45
	(93.83)	(93.83)	(152.75)	(311.97)	(88.59)	(88.59)	(143.05)	(334.25)	(1,634.70)
Net value of household transfers received, annualized	−7.44	−7.44	−6.50	−21.59	8.75	8.75	57.10*	7.45	130.18
	(13.06)	(13.06)	(15.69)	(32.14)	(19.10)	(19.10)	(30.66)	(58.05)	(263.75)
Tax paid, annualized	−0.09	−0.09	0.20	3.56	1.66	1.66	2.74	0.94	16.93
	(2.02)	(2.02)	(2.77)	(4.97)	(2.02)	(2.02)	(3.77)	(6.47)	(36.51)
Profits (ag & non-ag), annualized	33.73	33.73	90.18	183.24	44.08	44.08	91.17	125.89	485.20
	(48.95)	(48.95)	(84.27)	(149.09)	(45.35)	(45.35)	(77.81)	(154.16)	(787.10)
Wage earnings, annualized	73.72	73.72	96.93	201.44	182.99***	182.99***	180.03	−10.81	495.37
	(60.83)	(60.83)	(99.23)	(189.83)	(65.44)	(65.44)	(123.70)	(273.07)	(1,231.56)

Notes: Columns 1 and 5 replicate columns 2 and 3 from Table 1, selecting the number of radii bands included using our pre-specified algorithm as described in Section 3. The optimal radius selected is 2km for all outcomes. Columns 2-4 estimate the Total Effect (IV) for treated households, imposing a maximum radius R of 2, 4 and 6km respectively. Similarly, Columns 6-8 replicate Column 5, while including more spatial buffers. Standard errors are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.2: Robustness to fixing alternative radii bands: Input Prices and Quantities

	Recipient households				Non-recipient households				
	(1) Total Effect IV Optimal Radius	(2) Total Effect IV $\bar{R}=2$	(3) Total Effect IV $\bar{R}=4$	(4) Total Effect IV $\bar{R}=6$	(5) Total Effect IV Optimal Radius	(6) Total Effect IV $\bar{R}=2$	(7) Total Effect IV $\bar{R}=4$	(8) Total Effect IV $\bar{R}=6$	(9) Control, low-saturation mean (SD)
Panel A: Labor									
Hourly wage earned by employees	0.04	0.04	0.04	0.14	0.19*	0.19*	0.02	−0.11	0.70
	(0.04)	(0.04)	(0.07)	(0.10)	(0.10)	(0.10)	(0.16)	(0.30)	(0.89)
Household total hours worked, last 7 days	1.41	1.41	5.77	11.02	−4.70	−4.70	−3.16	−4.73	63.20
	(3.69)	(3.69)	(5.01)	(8.99)	(3.17)	(3.17)	(5.15)	(11.35)	(54.14)
Panel B: Land									
Land price per acre	365.44	365.44	694.99	815.15	556.83	556.83	556.44	−228.76	3,952.86
	(290.86)	(290.86)	(543.95)	(1,100.55)	(412.34)	(412.34)	(904.21)	(1,777.35)	(3,148.52)
Acres of land owned	−0.10	−0.10	0.02	0.10	0.08	0.08	0.12	0.04	1.42
	(0.09)	(0.09)	(0.16)	(0.55)	(0.10)	(0.10)	(0.17)	(0.40)	(2.37)
Panel C: Capital									
Loan-weighted interest rate, monthly	0.01	0.01	0.04**	0.02	−0.01	−0.01	−0.02	0.00	0.06
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.04)	(0.07)
Total loan amount	3.13	3.13	20.55*	59.32**	6.36	6.36	29.65	32.64	80.61
	(8.34)	(8.34)	(11.88)	(25.06)	(13.21)	(13.21)	(19.62)	(33.18)	(204.36)

Notes: Columns 1 and 5 replicate Columns 2 and 3 from Table 2, selecting the number of radii bands included using our pre-specified algorithm as described in Section 3. The optimal radius selected is 2km for all outcomes for both recipients and non-recipientsand groups. Columns 2, 3 and 4 estimate the Total Effect (IV) for treated households, imposing a maximum radius R of 2, 4 and 6km respectively. Similarly, Columns 6-8 replicate Column 5, while including more spatial buffers. Standard errors are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.3: Robustness to fixing alternative radii bands: Enterprise Outcomes

	Treatment Villages				Control Villages				
	(1) Total Effect IV Optimal Radius	(2) Total Effect IV $\bar{R}=2$	(3) Total Effect IV $\bar{R}=4$	(4) Total Effect IV $\bar{R}=6$	(5) Total Effect IV Optimal Radius	(6) Total Effect IV $\bar{R}=2$	(7) Total Effect IV $\bar{R}=4$	(8) Total Effect IV $\bar{R}=6$	(9) Control, low-saturation mean (SD)
Panel A: All enterprises									
Enterprise profits, annualized	67.53	67.53	156.38**	130.79	32.91	32.91	115.06	89.98	323.39
	(41.62)	(41.62)	(78.24)	(165.55)	(37.27)	(37.27)	(71.76)	(147.69)	(691.12)
Enterprise revenue, annualized	356.81**	356.81**	749.04***	831.13	244.27**	244.27**	596.71***	667.56	758.52
	(144.21)	(144.21)	(247.61)	(531.30)	(108.96)	(108.96)	(202.05)	(449.43)	(2,493.40)
Enterprise costs, annualized	96.82**	96.82**	93.50	−78.50	77.02	77.02	72.72	−97.34	147.73
	(40.92)	(40.92)	(63.04)	(145.28)	(48.95)	(48.95)	(68.60)	(145.39)	(550.11)
Enterprise wagebill, annualized	81.69**	81.69**	71.73	−55.13	70.49*	70.49*	63.91	−47.54	120.62
	(33.76)	(33.76)	(56.97)	(124.37)	(36.52)	(36.52)	(52.40)	(104.56)	(492.11)
Enterprise profit margin	−0.06**	−0.06**	−0.10*	−0.10	−0.06***	−0.06***	−0.10**	−0.11	0.44
	(0.03)	(0.03)	(0.05)	(0.09)	(0.02)	(0.02)	(0.05)	(0.09)	(0.61)
Panel B: Non-agricultural enterprises									
Enterprise inventory	34.68**	34.68**	33.39	−41.12	16.91	16.91	16.09	−49.24	192.98
	(14.73)	(14.73)	(20.68)	(48.61)	(10.80)	(10.80)	(14.79)	(41.76)	(504.76)
Enterprise investment, annualized	13.58	13.58	9.46	−22.43	6.82	6.82	3.22	−24.52	178.25
	(15.39)	(15.39)	(26.77)	(50.12)	(8.65)	(8.65)	(18.59)	(41.39)	(640.98)
Panel C: Village-level									
Number of enterprises	0.02	0.02	0.02	−0.04	0.01	0.01	0.01	−0.05	1.12
	(0.01)	(0.01)	(0.02)	(0.04)	(0.01)	(0.01)	(0.02)	(0.04)	(0.14)

Notes: Columns 1 and 5 replicate Columns 2 and 3 from Table 3, selecting the number of radii bands included using our pre-specified algorithm as described in Section 3. The optimal radius selected is 2km for all outcomes and groups. Columns 2, 3 and 4 estimate the Total Effect (IV) for enterprises in treated villages, imposing a maximum radius R of 2, 4 and 6km respectively. Similarly, Columns 5 and 6 replicate Column 4, while including more spatial buffers. Standard errors are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.4: BIC split sample approach for household expenditure, savings and income outcomes

	Recipient Households				Non-Recipient Households			
	Main Estimate		200 Split Sets		Main Estimate		200 Split Sets	
	(1) Total Effect (IV)	(2) Selected Radius	(3) Share in 95% CI	(4) Share selecting same radius	(5) Total Effect (IV)	(6) Selected Radius	(7) Share in 95% CI	(8) Share selecting same radius
Panel A: Expenditure								
Household expenditure, annualized	338.16***	2km	100%	100%	333.73***	2km	100%	100%
	(109.36)				(123.24)			
Non-durable expenditure, annualized	226.74**	2km	100%	100%	316.62***	2km	100%	100%
	(99.62)				(119.79)			
Food expenditure, annualized	133.55**	2km	100%	100%	132.84**	2km	100%	100%
	(63.98)				(58.58)			
Temptation goods expenditure, annualized	5.88	2km	100%	100%	−0.71	2km	100%	100%
	(8.82)				(6.50)			
Durable expenditure, annualized	109.07***	2km	100%	100%	8.41	2km	100%	100%
	(20.23)				(12.50)			
Panel B: Assets								
Assets (non-land, non-house), net borrowing	182.01***	2km	100%	99%	132.63*	2km	100%	100%
	(44.25)				(78.32)			
Housing value	480.68***	2km	100%	89%	72.58	2km	100%	100%
	(38.88)				(215.70)			
Land value	153.09	2km	100%	100%	572.07	2km	100%	99%
	(262.48)				(458.28)			
Panel C: Household balance sheet								
Household income, annualized	134.02	2km	100%	100%	229.46***	2km	100%	98%
	(93.83)				(88.59)			
Net value of household transfers received, annualized	−7.44	2km	100%	100%	8.75	2km	100%	100%
	(13.06)				(19.10)			
Tax paid, annualized	−0.09	2km	100%	100%	1.66	2km	100%	100%
	(2.02)				(2.02)			
Profits (ag & non-ag), annualized	33.73	2km	100%	100%	44.08	2km	100%	100%
	(48.95)				(45.35)			
Wage earnings, annualized	73.72	2km	100%	100%	182.99***	2km	100%	94%
	(60.83)				(65.44)			

Notes: Columns 1 and 2 reproduce the total effect estimates and radii selection for recipient households in Table 1. Columns 3 and 4 show the fraction of 200 random 50-50 splits of the sample for which the mean point estimates (across the two estimation samples within each split) for recipients falling into the 95%-CI of the initial estimate (Column 3) and select the same radius as the initial selection (Column 4). Columns 5 - 8 do the same for non-recipient households. See Table 1 for more details on variable construction and regression specification. Standard errors in Columns 1 and 5 are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.5: BIC split sample approach for input prices and quantities

	Recipient Households				Non-Recipient Households			
	Main Estimate		200 Split Sets		Main Estimate		200 Split Sets	
	(1) Total Effect (IV)	(2) Selected Radius	(3) Share in 95% CI	(4) Share selecting same radius	(5) Total Effect (IV)	(6) Selected Radius	(7) Share in 95% CI	(8) Share selecting same radius
Panel A: Labor								
Hourly wage earned by employees	0.04	2km	100%	100%	0.19*	2km	100%	93%
	(0.04)				(0.10)			
Household total hours worked, last 7 days	1.41	2km	100%	100%	−4.70	2km	100%	100%
	(3.69)				(3.17)			
Panel B: Land								
Land price per acre	365.44	2km	100%	99%	556.83	2km	100%	95%
	(290.86)				(412.34)			
Acres of land owned	−0.10	2km	100%	99%	0.08	2km	100%	100%
	(0.09)				(0.10)			
Panel C: Capital								
Loan-weighted interest rate, monthly	0.01	2km	99%	100%	−0.01	2km	100%	97%
	(0.01)				(0.01)			
Total loan amount	3.13	2km	100%	100%	6.36	2km	100%	100%
	(8.34)				(13.21)			

Notes: Columns 1 and 2 reproduce the total effect estimates and radii selection for recipient households in Table 2. Columns 3 and 4 show the fraction of 200 random 50-50 splits of the sample for which the mean point estimates (across the two estimation samples within each split) for recipients falling into the 95%-CI of the initial estimate (Column 3) and select the same radius as the initial selection (Column 4). Columns 5 - 8 do the same for non-recipient households. See Table 2 for more details on variable construction and regression specification. Standard errors in Columns 1 and 5 are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.6: BIC split sample approach for enterprise outcomes

	Treatment Villages		Control Villages		Radii Selected	
	(1) Total Effect (IV)	(2) Share in 95% CI	(3) Total Effect (IV)	(4) Share in 95% CI	(5) Main Estimate	(6) Share selecting same radius
Panel A: All enterprises						
Enterprise profits, annualized	67.53	100%	32.91	100%	2km	97%
	(41.62)		(37.27)			
Enterprise revenue, annualized	356.81**	100%	244.27**	100%	2km	97%
	(144.21)		(108.96)			
Enterprise costs, annualized	96.82**	90%	77.02	96%	2km	85%
	(40.92)		(48.95)			
Enterprise wagebill, annualized	81.69**	83%	70.49*	89%	2km	80%
	(33.76)		(36.52)			
Enterprise profit margin	−0.06**	100%	−0.06***	100%	2km	98%
	(0.03)		(0.02)			
Panel B: Non-agricultural enterprises						
Enterprise inventory	34.68**	99%	16.91	100%	2km	99%
	(14.73)		(10.80)			
Enterprise investment, annualized	13.58	100%	6.82	100%	2km	99%
	(15.39)		(8.65)			
Panel C: Village-level						
Number of enterprises	0.02	92%	0.01	92%	2km	94%
	(0.01)		(0.01)			

Notes: Columns 1 and 3 reproduce the total effect estimates for enterprises located in treatment and control villages from Table 3. Column 5 reports the radii selection (which is done across all enterprises jointly, since enterprises are not direct recipients of cash transfers). Columns 2 and 4 show the fraction of 200 random 50-50 splits of the sample for which the mean point estimates (across the two estimation samples within each split) produce estimates falling into the 95%-CI of the initial estimates in Columns 1 and 3 respectively. Column 6 shows the share of these splits where the algorighm selects the same radius as the initial selection in Column 5. See Table 3 for more details on variable construction and regression specification. Standard errors in Columns 1 and 3 are calculated following Conley (2008) using a uniform kernel out to 10 km. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table F.7: Maximum Radius Chosen by the BIC Algorithm (in km), expenditure, saving and income outcomes

	Recipient Households						Non-recipient Households					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full Sample	Low market access	High market access	Alego	Ugunja	Ukwala	Full Sample	Low market access	High market access	Alego	Ugunja	Ukwala
Panel A: Expenditure												
Household expenditure, annualized	2	2	2	2	2	2	2	2	2	2	2	2
Panel B: Assets												
Assets (non-land, non-house), net borrowing	2	2	2	2	2	2	2	2	2	2	2	2
Housing value	2	2	2	2	**4**	**4**	2	2	2	2	2	2
Land value	2	2	2	2	2	2	2	2	2	**4**	2	2
Panel C: Household balance sheet												
Household income, annualized	2	2	2	2	2	2	2	2	2	2	2	2
Net value of household transfers received, annualized	2	2	2	2	2	2	2	2	2	2	2	2
Tax paid, annualized	2	2	2	2	2	2	2	2	2	2	2	2
Profits (ag & non-ag), annualized	2	2	2	2	2	2	2	2	2	2	2	2
Wage earnings, annualized	2	2	2	2	2	2	2	2	2	2	2	2

Notes: This table reports the maximum radius selected minimizing a Baysian Information Criterion within different sub-samples of our data for outcomes in Table 1. Columns 1-6 report the radius selected when estimating effects on recipient households and when using the full sample (Column 1), when restricting the sample to villages with higher and lower than median market access (Columns 2 and 3), and to households in the three sub-counties Alego, Ugunja and Ukwala (Columns 4 - 6). Columns 7-12 report the same for non-recipient households.

Table F.8: Maximum Radius Chosen by the BIC Algorithm (in km), input prices and quantities

	Recipient Households						Non-recipient Households					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full Sample	Low market access	High market access	Alego	Ugunja	Ukwala	Full Sample	Low market access	High market access	Alego	Ugunja	Ukwala
Panel A: Labor												
Hourly wage earned by employees	2	2	2	2	**4**	2	2	2	2	2	2	2
Household total hours worked, last 7 days	2	2	2	2	2	2	2	2	2	2	2	2
Panel B: Land												
Land price per acre	2	2	2	2	**16**	2	2	2	2	2	2	2
Acres of land owned	2	2	2	2	2	2	2	2	2	2	2	2
Panel C: Capital												
Loan-weighted interest rate, monthly	2	2	2	2	2	2	2	2	2	2	2	**4**
Total loan amount	2	2	2	2	2	2	2	2	2	2	2	2

Notes: This table reports the maximum radius selected minimizing a Baysian Information Criterion within different sub-samples of our data for outcomes in Table 2. Columns 1-6 report the radius selected when estimating effects on recipient households and when using the full sample (Column 1), when restricting the sample to villages with higher and lower than median market access (Columns 2 and 3), and to households in the three sub-counties Alego, Ugunja and Ukwala (Columns 4 - 6). Columns 7-12 report the same for non-recipient households.

Table F.9: Maximum Radius Chosen by the BIC Algorithm (in km), enterprise outcomes

		Market Access		**Subcounty**		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	Low market access	High market access	Alego	Ugunja	Ukwala
Panel A: All enterprises						
Enterprise profits, annualized	2	**4**	2	2	2	2
Enterprise revenue, annualized	2	**4**	2	2	2	2
Enterprise costs, annualized	2	**8**	2	2	2	2
Enterprise wagebill, annualized	2	2	2	2	2	2
Enterprise profit margin	2	2	2	2	2	2
Panel B: Non-agricultural enterprises						
Enterprise inventory	2	**8**	2	2	2	2
Enterprise investment, annualized	2	2	2	2	2	2
Panel C: Village-level						
Number of enterprises	2	2	2	2	2	2

Notes: This table reports the maximal radius selected by minimizing a Baysian Information Criterion within different sub-samples of our data for enterprise outcomes in Table 3. We report the radius selected when using the full sample (Column 1), when restricting the sample to villages with higher and lower than median market access (Columns 2 and 3), and to enterprises in the three sub-counties Alego, Ugunja and Ukwala (Columns 4 - 6).

Table F.10: Randomization inference for expenditure, savings and income outcomes

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	Total Effect IV	Spatial RI p-value	Total Effect IV	Spatial RI p-value
Panel A: Expenditure				
Household expenditure, annualized	338.16***	[0.00]***	333.73***	[0.02]**
	(109.36)		(123.24)	
Non-durable expenditure, annualized	226.74**	[0.03]**	316.62***	[0.02]**
	(99.62)		(119.79)	
Food expenditure, annualized	133.55**	[0.07]*	132.84**	[0.07]*
	(63.98)		(58.58)	
Temptation goods expenditure, annualized	5.88	[0.59]	−0.71	[0.90]
	(8.82)		(6.50)	
Durable expenditure, annualized	109.07***	[0.00]***	8.41	[0.43]
	(20.23)		(12.50)	
Panel B: Assets				
Assets (non-land, non-house), net borrowing	182.01***	[0.00]***	132.63*	[0.24]
	(44.25)		(78.32)	
Housing value	480.68***	[0.00]***	72.58	[0.80]
	(38.88)		(215.70)	
Land value	153.09	[0.61]	572.07	[0.29]
	(262.48)		(458.28)	
Panel C: Household balance sheet				
Household income, annualized	134.02	[0.16]	229.46***	[0.07]*
	(93.83)		(88.59)	
Net value of household transfers received, annualized	−7.44	[0.51]	8.75	[0.59]
	(13.06)		(19.10)	
Tax paid, annualized	−0.09	[0.97]	1.66	[0.43]
	(2.02)		(2.02)	
Profits (ag & non-ag), annualized	33.73	[0.46]	44.08	[0.47]
	(48.95)		(45.35)	
Wage earnings, annualized	73.72	[0.36]	182.99***	[0.06]*
	(60.83)		(65.44)	

Notes: This table presents randomization inference results for outcomes in Table 1. Column 1 reproduces the total effect for recipient households (Column 2 in Table 1), and Column 3 reproduces the total effect for non-recipient households (Column 3 in Table 1). Columns 2 and 4 report randomization inference p-values from 500 repetitions (i.e. the share of replications where the estimated total effect is larger than the main estimate), where the cross-sectional distribution of cash distributed in radii bands across villages was randomly shuffled.

Table F.11: Randomization inference for input prices and quantities

	(1)	(2)	(3)	(4)
	Recipient Households		**Non-recipient Households**	
	Total Effect IV	Spatial RI p-value	Total Effect IV	Spatial RI p-value
Panel A: Labor				
Hourly wage earned by employees	0.04	[0.55]	0.19*	[0.21]
	(0.04)		(0.10)	
Household total hours worked, last 7 days	1.41	[0.73]	−4.70	[0.15]
	(3.69)		(3.17)	
Panel B: Land				
Land price per acre	365.44	[0.33]	556.83	[0.14]
	(290.86)		(412.34)	
Acres of land owned	−0.10	[0.64]	0.08	[0.56]
	(0.09)		(0.10)	
Panel C: Capital				
Loan-weighted interest rate, monthly	0.01	[0.62]	−0.01	[0.62]
	(0.01)		(0.01)	
Total loan amount	3.13	[0.77]	6.36	[0.68]
	(8.34)		(13.21)	

Notes: This table presents randomization inference results for outcomes in Table 2. Column 1 reproduces the total effect for recipient households (Column 2 in Table 2), and Column 3 reproduces the total effect for non-recipient households (Column 3 in Table 2). Columns 2 and 4 report randomization inference p-values from 500 repetitions (i.e. the share of replications where the estimated total effect is larger than the main estimate), where the cross-sectional distribution of cash distributed in radii bands across villages was randomly shuffled.

Table F.12: Randomization inference for enterprise outcomes

	(1)	(2)	(3)	(4)
	Treated Villages		**Control Villages**	
	Total Effect IV	Spatial RI p-value	Total Effect IV	Spatial RI p-value
Panel A: All enterprises				
Enterprise profits, annualized	67.53	[0.23]	32.91	[0.43]
	(41.62)		(37.27)	
Enterprise revenue, annualized	356.81**	[0.05]**	244.27**	[0.06]*
	(144.21)		(108.96)	
Enterprise costs, annualized	96.82**	[0.03]**	77.02	[0.03]**
	(40.92)		(48.95)	
Enterprise wagebill, annualized	81.69**	[0.04]**	70.49*	[0.03]**
	(33.76)		(36.52)	
Enterprise profit margin	−0.06**	[0.12]	−0.06***	[0.03]**
	(0.03)		(0.02)	
Panel B: Non-agricultural enterprises				
Enterprise inventory	34.68**	[0.03]**	16.91	[0.12]
	(14.73)		(10.80)	
Enterprise investment, annualized	13.58	[0.33]	6.82	[0.47]
	(15.39)		(8.65)	
Panel C: Village-level				
Number of enterprises	0.02	[0.23]	0.01	[0.56]
	(0.01)		(0.01)	

Notes: This table presents randomization inference results for outcomes in Table 3. Column 1 reproduces the total effect for enterprises in treated villages (Column 2 in Table 3), and Column 3 reproduces the total effect for enterprises in control villages (Column 3 in Table 3). Columns 2 and 4 report randomization inference p-values from 500 repetitions (i.e. the share of replications where the estimated total effect is larger than the main estimate), where the cross-sectional distribution of cash distributed in radii bands across villages was randomly shuffled.

Table F.13: Randomization inference for price outcomes

		(1)	(2)	(3)	(4)
		Overall Effects		**ATE by market access**	
		ATE	Average maximum effect (AME)	below median	above median
All goods		0.0010*	0.0042	0.0017*	0.0007
		(0.0006)	(0.0031)	(0.0009)	(0.0007)
		[0.070]	[0.110]	[0.098]	[0.500]
By tradability	More tradable	0.0014	0.0062	0.0023	0.0021
		(0.0015)	(0.0082)	(0.0023)	(0.0018)
		[0.268]	[0.370]	[0.264]	[0.382]
	Less tradable	0.0009	0.0034	0.0015	0.0001
		(0.0006)	(0.0032)	(0.0011)	(0.0008)
		[0.268]	[0.460]	[0.536]	[0.856]
By sector	Food items	0.0009	0.0036	0.0016	0.0002
		(0.0006)	(0.0033)	(0.0012)	(0.0008)
		[0.248]	[0.434]	[0.528]	[0.796]
	Non-durables	0.0014	0.0061	0.0026	0.0019
		(0.0017)	(0.0089)	(0.0026)	(0.0019)
		[0.310]	[0.422]	[0.234]	[0.416]
	Durables	0.0019*	0.0070	-0.0009	0.0034**
		(0.0011)	(0.0061)	(0.0011)	(0.0016)
		[0.052]	[0.152]	[0.630]	[0.094]
	Livestock	-0.0008	-0.0027	-0.0008*	-0.0017
		(0.0010)	(0.0052)	(0.0004)	(0.0020)
		[0.224]	[0.534]	[0.288]	[0.500]
	Temptation goods	-0.0011	-0.0112	-0.0008	-0.0003
		(0.0026)	(0.0143)	(0.0036)	(0.0035)
		[0.734]	[0.522]	[0.810]	[0.946]

Notes: This table presents randomization inference results for outcomes in Table 4. Randomization inference p-values from 500 repetitions (i.e. the share of replications where the estimated total effect is larger than the main estimate) are reported in brackets. In each iteration, we randomly re-generate cross-sectional and temporal roll out of transfers according to our actual treatment randomization as described in Section 2.2.

G Estimating the marginal propensity to consume and spend locally

This appendix section provides details on the marginal propensity to consume (MPC) estimates reported in section 4.2.

We focus on recipients' propensity to spend as a result of the transfer, which is directly relevant for the first-round spending impacts in the local economy. To the extent that recipient households generate additional earned income as the result of the transfer, and also spend out of this income, the main marginal propensity to consume estimate may be an overestimate. Below, we therefore also present recipient expenditure effects relative to the transfer amount received plus any additional income generated as a result of the transfer. (We are also able to obtain an estimate of the marginal propensity to consume among transfer non-recipients, by taking the ratio of spending impacts relative to income effects over the same time period. In fact, the estimates for non-recipients are quantitatively similar to those estimated among cash transfer recipients. Since income is likely to be imperfectly measured relative to expenditure in this context, see Deaton and Zaidi (2002), and because first-round spending impacts are particularly important, we focus on the MPC among transfer recipients.)

In rural African settings like ours, formal sector financial savings (e.g., in bank accounts) or cash savings are limited. Only 10% of households in our study area report having a bank account at endline. In ongoing work in a similar Kenyan context, total savings in mobile money, cash and bank accounts amounted to roughly 100 USD PPP in the control group, a small share of total assets. The effect on total savings of a 1000 USD PPP transfer (which is roughly half the size of the transfer in our study) after 14 months was only roughly 25 USD PPP, or 2.5% of the transfer. Instead, most household saving comes in the form of purchases of relatively liquid durable assets such as livestock or even housing materials. In what follows, we separately present recipient spending on durable assets and non-durable consumption goods. From an intertemporal decision-making perspective, the latter represents pure "consumption", while the former is likely have both a "consumption" and a "savings" component.

Whether they are "consumed" or "saved", expenditures on both durables and non-durables are predominantly local: over 95% of respondents report shopping locally for both types of goods. In a context where financial savings options are limited, high marginal propensities to spend — which as noted above, is not necessarily the same as to consume — should not be unexpected. From the perspective of quantifying the transfer multiplier, it is this marginal propensity to *spend* that matters, as spending on both "consumption goods" and "savings goods" show up as revenue for local firms, and have a similar stimulus effect on the local economy. Our main estimate of the MPC (MPC total) therefore includes both components.

Importantly, recipient expenditures only enter the local economy, and thus generate a local multiplier, if they occur locally and contribute to the income of another local agent. We call the measure of this type of expenditure the marginal propensity to spend locally (MPC local). Since the vast majority of individuals in the study sample work locally and firms are overwhelmingly locally owned (as noted in the main text), we expect nearly all factor payments to remain in the local economy. The main reason why local revenue might not end up as local income is the importing of intermediate goods. In Appendix Section

H.1, we calculate that up to 19% of non-durable consumption and 20% of durable purchases indirectly reflect imports of intermediate goods from outside the study area. Our preferred measure of the MPC local adjusts the overall marginal propensity to spend (MPC total) to account for such imports, leading the MPC local to be smaller in magnitude than MPC total.

Table G.1: Estimates of recipients' marginal propensity to consume

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Transfer					**Transfer + Income Gains**	
	MPC non-durables		**MPC durables**	**MPC total**	**MPC local**	**MPC total**	**MPC local**
	q1-q3	q4-q10					
Our data only	-0.22	0.29	0.30	0.37	0.30	0.33	0.27
	(0.22)	(0.12)	(0.05)	(0.24)	(0.19)	(0.19)	(0.16)
Rarieda data q1-3, our data q4-10	0.35	0.29	0.30	0.93	0.76	0.84	0.68
	(0.11)	(0.12)	(0.05)	(0.16)	(0.13)	(0.13)	(0.10)

Notes: This table presents estimates of recipients' marginal propensity to consume. Columns 1 and 2 report total effects on non-durable expenditure over the first 3 and the next 7 quarters after the transfer respectively relative to the average treatment amount received, and estimated dynamically according to Equation 7. Column 3 adds the effect on accumulated durable assets (including house value) for recipients at endline, estimated using Equation 2. Column 4 sums up Columns 1-3, presenting our main marginal propensity to spend (MPC total). Column 5 adjusts Column 4 by accounting for an estimated 20% of durables and 18% of non-durables expenditure that reflects imports of intermediates, yielding the marginal propensity to spend locally (MPC local). See Appendix H.1 for details. Columns 6 and 7 present these effects relative to the average transfer amount plus the average additional income recipients' generated over the 27 months after the transfer, again estimated using Equation 2. The first row estimates the MPC using only data from this study. The second row estimates the first 3 quarters of the non-durable expenditure effect in Column 1 using midline expenditure data from households in neighboring Rarieda county that received similar transfers as part of Haushofer and Shapiro (2016), which has the advantage of capturing the immediate expenditure response to the transfer. It is estimated analogously using Equation 7, but excluding spillover terms. Standard errors (in parentheses) come from 2000 iterations of a wild-bootstrap, clustered at the sublocation level for our data, and the village level for Rarieda data.

Row 1 in Table G.1 presents MPC estimates using data only from this study. We estimate a marginal propensity to spend on non-durables of 0.07 over the first 10 quarters after the initial transfer, and 0.30 on on durables. Combined, this yields a marginal propensity to spend of 0.37. We are thus able to directly account for 37% of the transfer. Adjusted for imported intermediates, the MPC local is 0.30.

As noted in the main text, this estimate faces the important limitation that the endline data collection started about 9 months after transfers in a village went out (see FigureA.1b). Recall periods for non-durable consumption range from a week to a month, making it hard to convincingly estimate *direct* spending effects for recipients on non-durables in the initial months after the transfer. Regarding measures from our data, we show in Figure 1 that the observed variation in transfer amounts received in the 3 quarters prior to our surveys is limited, leading to imprecise estimates (that are also small or even negative). This limitation is less relevant for the estimation of across-village spillovers because transfers to surrounding villages may have randomly gone out earlier or later relative to the timing of the survey, thus providing ample variation to estimate early spillover effects over a short time horizon. Note that estimates of durables expenditure effects do not face the same problem as non-durable consumption, since we measure these as the difference in the stock of durables at endline (between treatment and control), rather than as an integral of flows.

To overcome this limitation, the second row of Table G.1 brings in additional evidence from a closely related project in the neighboring sub-county of Rarieda that collected more detailed data on household spending in the months immediately after receipt of similar GiveDirectly transfers (Haushofer and Shapiro, in prep.). Specifically, we estimate the direct impact of transfers on recipient spending, estimating the dynamic impact according to Equation 7 as we do for our data, but excluding spillover terms (due to a different spatial randomization strategy in Rarieda); as we expect spillovers to be less important relative to direct effects in the initial months, this is a relatively mild assumption.

The data in Column 1 shows that initial direct spending impacts on non-durable goods in Rarieda were indeed far higher than what we estimate in our data, at 0.35. Combined with our data on non-durable expenditure in the quarters thereafter, we estimate that recipients' spend 64% of the transfer on non-durables over the first 10 quarters. Adding in durable expenditure yields our preferred estimate of the marginal propensity to spend (MPC total) of 0.93. This indicates that we are close to accounting for the entire transfer amount being spent, and highlights that the study population can be characterized as largely hand-to-mouth consumers. Even when we account for increased income generated by recipients over the same period in Column 6, the estimate of the total marginal propensity to spend remains very high, at 0.84. This is again in line with the observation that savings in formal financial products or even in cash are unlikely to be substantial in this context.

The preferred estimate of the marginal propensity spend locally, which accounts for imports of intermediate goods is presented in Column 5, and yields an estimated MPC local of 0.76. An alternative estimate that accounts for any additional income generated (among transfer recipients) is similar, at 0.68 (Column 7). These calculations illustrate that a large share of transfer is spent by recipient households within our study period, and roughly three quarters re-enters the local economy and ends up as income of another local agent. In a simple static Keynesian framework, an MPC local in the range of 0.68 to 0.76 implies a local economy transfer multiplier $\frac{MPC}{1-MPC}$ between 2.1 to 3.2

H Transfer multiplier - robustness

This section conducts three main robustness checks regarding the multiplier analysis. In the first subsection, we attempt to account for transactions between agents in our study area and those located outside it. Using a combination of household and enterprise data, and conservative assumptions on import shares by type of enterprise, we provide an upper bound on the share of the expenditure multiplier that may reflect increased imports from outside the study area. The second subsection makes alternative assumptions about the expenditure effects in the initial months after transfers, which as noted in the main text are noisily estimated in our data because the average endline survey took place 18 months after the first transfers were received. Third, we present estimates in nominal terms (rather than real terms).

H.1 Accounting for imports of intermediate goods

As described in Section 5, the main expenditure multiplier incorrectly includes imports which are not part of local value added. There are many reasons to believe that any resulting bias is relatively small. From household shopping patterns, we know that only 10% of households report ever shopping at a market outside our study area. Non-farm businesses report only 5% of customers coming from outside the study area. In addition, the estimated effects on household consumption and enterprise revenue are fairly similar, suggesting that consumer spending was quite localized and direct imports by households are relatively small. The main concern is therefore imported intermediate goods.

To gauge whether this bias is quantitatively important, we first assign each component of our non-durable expenditure and durable asset measures to one of 48 enterprise types where it is most likely to be purchased. When there are multiple possible types of enterprises, we use overall revenue shares of different enterprise types to distribute expenditure between them. Reassuringly, this correspondence implies expenditure shares by enterprise type that match their revenue shares from the enterprise survey fairly well (correlation coefficient of 0.62). For each enterprise type, we then obtain an upper-bound for the share of intermediate inputs in overall value added as: $1 - \frac{1}{N}\sum_i w_i \frac{cost_i + profit_i}{revenue_i}$ (where we first winsorize at the 1% and 99% cut-offs, then average across enterprises of each type using revenue weights, and cap at 0 and 1), and where i denotes a firm and N is the total number of firms of that type, and w_i the revenue weight of firm i (re-scaled to sum to 1). This is clearly an upper bound, since the enterprise survey cost measure only contains selected components of firm costs.

Next, we make assumptions based on an understanding of the local context about what share of intermediate inputs is imported from outside the study area. In doing so, we try to err on the side of an import share that is too high. The total share of imports in consumption expenditure and assets is then calculated as the expenditure-weighted share of imports of intermediate goods for each expenditure and asset category. For the exact correspondence between each consumption good or asset and enterprise types, consult Tables H.1 and H.2.

Using this methodology, the upper bound estimate of the share of imports in non-durable consumption goods is 18%, and for assets, the figure is 20%. This shows that imports of intermediate goods may be non-negligible, but that a large majority of spending still reflects local economic activity (and recall that these figures are upper bounds). To get a sense

Table H.1: Non-durable expenditure: Intermediate input and import shares

Item	Bought atenterprise type	(1) Expenditure share (data)	(2) Intermediate input share (data)	(3) Intermediate import share (assumed)	(4) Overall import share
Cereals	Cereals	5.9%	60%	50%	30%
	Posho mill	5.9%	26%	0%	0%
	Small retail	2.9%	65%	75%	49%
Roots and tubers	Food stall / Raw food and fruits vendor	2.6%	44%	25%	11%
Pulses	Food stall / Raw food and fruits vendor	3.7%	44%	25%	11%
Vegetables	Food stall / Raw food and fruits vendor	8.6%	44%	25%	11%
Fruits	Food stall / Raw food and fruits vendor	2.9%	44%	25%	11%
Meat	Butcher	4.2%	58%	0%	0%
	Livestock / Animal (Products) / Poultry Sale	0.5%	20%	50%	10%
Fish	Fish Sale / Mongering	6.0%	41%	0%	0%
Dairy and eggs	Food stall / Raw food and fruits vendor	4.6%	44%	25%	11%
Other animal products	Livestock / Animal (Products) / Poultry Sale	0.5%	20%	50%	10%
Cooking fat	Small retail	3.7%	65%	75%	49%
Sugar products	Jaggery	2.6%	54%	0%	0%
	Small retail	2.6%	65%	75%	49%
Jam, honey, sweets, candies	Small retail	0.2%	65%	75%	49%
Tea, coffee	Small retail	1.5%	65%	75%	49%
Salt, pepper, condiments, etc.	Small retail	0.7%	65%	75%	49%
Food eaten outside the house	Food stand / Prepared food vendor	0.8%	56%	25%	14%
	Restaurant	0.6%	48%	50%	24%
Alcohol, tobacco	Bar	0.2%	41%	100%	41%
	Homemade alcohol / liquor	1.0%	52%	0%	0%
	Small retail	0.5%	65%	75%	49%
Other foods	Small retail	0.3%	65%	75%	49%
Clothing and shoes	Clothes / Mtumba / Boutique	1.0%	37%	100%	37%
	Tailor	1.8%	18%	100%	18%
Personal items	Barber shop	0.8%	0%	100%	0%
	Beauty shop / Salon	0.2%	12%	100%	12%
	Photo studio	0.0%	0%	100%	0%
	Small retail	1.0%	65%	75%	49%
Household items	Small retail	2.3%	65%	75%	49%
Transport, travel	Guesthouse/ Hotel	0.5%	18%	75%	14%
	Petrol station	2.3%	86%	100%	86%
	Piki driver	1.9%	26%	100%	26%
Airtime and phone expenses	M-Pesa	2.7%	54%	100%	54%
Internet	Cyber café	0.1%	18%	100%	18%
Firewood, charcoal, kerosene	Charcoal sale / burning	1.6%	16%	0%	0%
	Kerosene	0.1%	36%	100%	36%
	Timber / Firewood	0.1%	45%	50%	22%
Electricity	Local	0.3%		0%	0%
Water	Local	0.3%		0%	0%
Recreation	Bookshop	0.0%	21%	100%	21%
	Small retail	0.1%	65%	75%	49%
	Video Room/Football hall	0.0%	57%	100%	57%
Lottery tickets and gambling	Small retail	0.1%	65%	75%	49%
Religious expenses	Local	0.6%		0%	0%
Weddings, funerals	Local	1.0%		0%	0%
Charitable expenses	Local	0.1%		0%	0%
House rent / mortgage	Local	0.5%		0%	0%
School expenses	Local	10.7%		0%	0%
Medical expenses	Chemist	2.3%	27%	100%	27%
Other expenses	Local	4.2%		0%	0%
Total	**100.0%**				**18%**

Notes: Each row corresponds to an item in the expenditure module of our household surveys. We match each expenditure item to the enterprise type (from our enterprise census) at which it was most likely purchased (using revenue shares where possible to distribute expenditure where a good may be purchased in multiple enterprise types). Column 1 contains expenditure shares (sample-weighted across all households). Column 2 reports the upper bound of the share of intermediates in value added for each enterprise type, calculated from our enterprise surveys as $1 - \frac{1}{N}\sum_i w_i \frac{cost_i + profit_i}{revenue_i}$ (where we first winsorize at the 1% and 99% cut-offs, then average across enterprises of each type using revenue weights w_i, and cap at 0 and 1). Column 3 contains our upper-bound assumptions about the share of intermediate goods for each enterprise type that is imported, and Column 4 displays the implied import share for this item (multiplying Columns 2 and 3). The last row is expenditure-weighted across all categories.

Table H.2: Durable assets: Intermediate input and import shares

Item	Bought at enterprise type	(1) Asset share (data)	(2) Intermediate input share (data)	(3) Intermediate import share (assumed)	(4) Overall import share
Bicycle	Import	0.5%		100%	100%
	Bicycle repair / mechanic shop	0.5%	0%	100%	0%
Motorcycle	Motorcycle Repair / Shop	0.5%	45%	100%	45%
	Import	1.9%		100%	100%
Car	Import	2.5%		100%	100%
Boat	Import	0.0%		100%	100%
Bed	Carpenter	2.0%	10%	75%	7%
Chair	Carpenter	1.1%	10%	75%	7%
Table	Carpenter	1.3%	10%	75%	7%
Cupboard	Carpenter	1.5%	10%	75%	7%
Sofa	Carpenter	4.2%	10%	75%	7%
Mattress	Import	1.8%		100%	100%
Bednet	Hardware store	0.1%	41%	100%	41%
Solar energy system	Electric accesory/repair	0.3%	6%	100%	6%
	Import	1.0%		100%	100%
Generator	Hardware store	0.1%	41%	100%	41%
Car battery	Hardware store	0.2%	41%	100%	41%
Kerosene	Kerosene	0.1%	36%	100%	36%
Lantern	Hardware store	0.2%	41%	100%	41%
Clock	Electric accesory/repair	0.1%	6%	100%	6%
Radio	Electric accesory/repair	0.6%	6%	100%	6%
Sewing machine	Electric accesory/repair	0.4%	6%	100%	6%
Electric Iron	Electric accesory/repair	0.0%	6%	100%	6%
Mobile phone	Electric accesory/repair	0.7%	6%	100%	6%
	Import	0.7%		100%	100%
Television	Electric accesory/repair	0.7%	6%	100%	6%
Computer	Electric accesory/repair	0.0%	6%	100%	6%
	Import	0.0%		100%	100%
Cattle	Livestock / Animal (Products) / Poultry Sale	11.4%	20%	50%	10%
Pig	Livestock / Animal (Products) / Poultry Sale	0.3%	20%	50%	10%
Sheep	Livestock / Animal (Products) / Poultry Sale	0.6%	20%	50%	10%
Goat	Livestock / Animal (Products) / Poultry Sale	0.6%	20%	50%	10%
Chicken	Livestock / Animal (Products) / Poultry Sale	1.4%	20%	50%	10%
Other birds	Livestock / Animal (Products) / Poultry Sale	0.1%	20%	50%	10%
Farm tools	Hardware store	0.6%	41%	100%	41%
Ox plow	Hardware store	0.1%	41%	100%	41%
Wheel barrow	Hardware store	0.3%	41%	100%	41%
Hand cart	Hardware store	0.0%	41%	100%	41%
Iron sheets	Hardware store	0.4%	41%	100%	41%
House value (maintenance, improvement)	Carpenter	12.3%	10%	75%	7%
	Welding / metalwork	12.3%	0%	100%	0%
	Hardware store	18.4%	41%	100%	41%
	Local	18.4%		0%	0%
Total	**100.0%**				**20%**

Notes: Each row corresponds to an item in the asset module of our household surveys. We match each asset to the enterprise type (from our enterprise census) at which it was most likely purchased (using revenue shares where possible to distribute assets where a good may be purchased in multiple enterprise types). Column 1 contains asset shares (sample-weighted across all households). Column 2 reports the upper bound of the share of intermediates in value added for each enterprise type, calculated from our enterprise surveys as $1 - \frac{1}{N}\sum_i w_i \frac{cost_i + profit_i}{revenue_i}$ (where we first winsorize at the 1% and 99% cut-offs, then average across enterprises of each type using revenue weights w_i, and cap at 0 and 1). Column 3 contains our upper-bound assumptions about the share of intermediate goods for each enterprise type that is imported, and Column 4 displays the implied import share for this item (multiplying Columns 2 and 3). The last row is asset-share-weighted across all categories.

of how this impacts the multiplier estimate, we further assume that (i) all inventories are in the form of intermediate goods rather than final goods (leading us to err on the side of overstating their import share, at 62%), (ii) the import share of enterprise investment is the same as that of household assets (in our context, household and enterprise assets are often comparable or even shared), and (iii) imports scale linearly with expenditure. We then compute the share of the expenditure-based multiplier that is spent locally (see Table H.3). Even under the set of conservative assumptions discussed above, the transfer multiplier for local expenditure remains similar at 2.01.

Table H.3: Transfer Multiplier Estimates: Adjusting for Imported Intermediates

	(1) $\mathbb{M}$ Estimate	(2) Share imported	(3) Import adjusted
Panel A: Expenditure multiplier	2.53	0.20	2.01
Household non-durable expenditure	1.17	0.18	0.96
Household durable expenditure	0.81	0.20	0.65
Enterprise investment	0.48	0.20	0.38
Enterprise inventory	0.07	0.62	0.03

Notes: Results from the joint estimation of the expenditure multiplier (as in Table 5). Column 1 reports our main point estimates of the expenditure multiplier components. Each component is estimated individually and the multiplier is obtained by aggregating components as described in the main text. Column 2 presents our upper-range estimates of the share of imports captures for each expenditure component, and Column 3 presents the import-share adjusted estimate of the multiplier on local expenditure only.

H.2 Alternative assumptions for initial expenditure responses

Figure A.1 illustrates the timing of endline household surveys and enterprise surveys, and the substantial time lag between when the first transfers were scheduled in each village and survey administration (with time lags of 9 and 18 months, respectively). This limitation implies that treatment effects on expenditures in the three quarters post-transfer are quite imprecisely estimated, as discussed in detail in Appendix Section G.

We conduct two additional robustness tests to partially address these concerns. First, from the perspective of minimizing mean squared error, it may be preferable to exclude the noisily estimated initial quarters for all components of each multiplier that rely on flow values. This almost certainly leads to a substantial downward bias, since all early spending, profits and investment are excluded, but may improve statistical precision. The results of this exercise are presented in Panel A of Table H.4. As expected, the estimated expenditure and income multipliers are both lower compared to the main specifications, with the average of both multipliers falling to 1.75. The standard standard error on this estimate also declines substantially, by more than half. When testing both multipliers jointly, we reject a multiplier smaller than one with a p-value of 0.04.

Second, we utilize data from a closely related project in a neighboring county Rarieda that collected more detailed data on recipient household spending in the months immediately after they received similar GiveDirectly transfers a few years prior to this experiment (Haushofer and Shapiro 2013). While this project did not collect data on non-recipients, its data complements our data precisely where we think the timing of surveys and transfers imposes the most significant limitation for us, namely for estimating the direct impacts of transfers on *recipients* in the initial period post-transfer. In this exercise, we replace the noisily estimated consumption impacts among recipient households in the first 3 quarters post-transfer with estimates from the Rarieda data. Specifically, we estimate the same equation 7 as we do for our data, but exclude spillover terms (see Appendix G for more details). For all other components, and for responses among non-recipients, the inputs into the multiplier estimate are unchanged.

Panel A of Table H.4 shows that augmenting the spending impact estimates with the data from Haushofer and Shapiro (2013) leads to a larger expenditure multiplier estimate of 3.09 (that is also slightly more precisely estimated than our main estimate). When testing both multipliers jointly, we reject a multiplier smaller than one with a p-value of 0.04.

Table H.4: Transfer Multiplier: Alternative Assumptions for the Initial Spending Impact

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Main estimate	**Alternative Specification I: Setting initial 3 quarters = 0**			**Alternative Specification II Initial 3 quarters from Haushofer & Shapiro (2016)**		
		$\mathbb{M}$ Estimate	H_0: $\mathbb{M} < 0$ *p*-value	H_0: $\mathbb{M} < 1$ *p*-value	$\mathbb{M}$ Estimate	H_0: $\mathbb{M} < 0$ *p*-value	H_0: $\mathbb{M} < 1$ *p*-value
Panel A: Expenditure multiplier	2.53 (1.42)	2.04 (0.67)	0.00***	0.06*	3.09 (1.38)	0.01**	0.06*
Household non-durable expenditure	1.17 (1.32)	0.99 (0.63)	0.06*		1.73 (1.25)	0.08*	
Household durable expenditure	0.81 (0.05)	0.81 (0.05)	0.00***		0.81 (0.05)	0.00***	
Enterprise investment	0.48 (0.42)	0.17 (0.11)	0.06*		0.48 (0.44)	0.15	
Enterprise inventory	0.07 (0.03)	0.07 (0.03)	0.02**		0.07 (0.03)	0.02**	
Panel B: Income multiplier	2.28 (1.73)	1.45 (0.65)	0.01***	0.23	2.28 (1.76)	0.12	0.24
Enterprise profits	1.47 (1.28)	0.00 (0.35)	0.48		1.47 (1.28)	0.14	
Household wage bill	0.68 (1.15)	1.34 (0.54)	0.01***		0.68 (1.15)	0.28	
Enterprise capital income	0.09 (0.17)	0.10 (0.06)	0.05*		0.09 (0.17)	0.32	
Enterprise taxes paid	0.04 (0.03)	0.01 (0.01)	0.03**		0.04 (0.03)	0.08*	
Panel C: Expenditure and income multipliers							
Average of both multipliers	2.40 (1.38)	1.75 (0.58)	0.00***	0.09*	2.69 (1.39)	0.03**	0.12
Joint test of both multipliers			0.00***	0.04**		0.01***	0.04**

Notes: Results from the joint estimation of expenditure and income multipliers. Column 1 reports our main point estimates of both multipliers and their respective components from Table 5. Columns 2 - 4 repeat this exercise, imposing that the impact of each dynamically estimated flow component is zero in the first 3 quarters after the transfer. Columns 5 - 7 estimate the initial 3 quarters of the impact on non-durable consumption expenditure for recipients using data from a related project that collected more detailed data for recipient expenditure in the initial months after the transfer (Haushofer and Shapiro 2013). All other components remain the same as in our main specification. Transfer amounts and outcome variables are deflated to January 2015 using the overall consumer price index in the geographically closest market. Standard errors are computed by 2,000 replications of a clustered wild clustered bootstrap. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

H.3 The nominal transfer multiplier

The main multiplier estimate is based on real GDP, in which transfer amounts and all outcome measures are deflated to January 2015 US Dollars using the overall consumer price index in the geographically closest market to each household or enterprise (see Section 3.6 for a description of the price data). Table H.5 presents the same exercise in nominal terms. Since we estimate small treatment effects on prices, the difference between the real and nominal measures is mainly driven by overall inflation in the study area. As shown in Figure E.1, prices in the study area rose by about 10% per year on average. Roughly in line with this, the nominal multiplier over the first two years after transfers went out is roughly 5% larger than the real multiplier (2.66 versus 2.53) on the expenditure side, and approximately 12% larger (2.55 versus 2.28) on the income side.

Table H.5: Nominal Transfer Multiplier

	(1) $\mathbb{M}$ Estimate	(2) H_0: $\mathbb{M} < 0$ p-value	(3) H_0: $\mathbb{M} < 1$ p-value
Panel A: Expenditure multiplier	2.66 (1.48)	0.04**	0.12
Household non-durable expenditure	1.22 (1.37)	0.18	
Household durable expenditure	0.89 (0.06)	0.00***	
Enterprise investment	0.47 (0.43)	0.15	
Enterprise inventory	0.08 (0.04)	0.02**	
Panel B: Income multiplier	2.55 (1.80)	0.08*	0.19
Enterprise profits	1.47 (1.30)	0.13	
Household wage bill	0.94 (1.17)	0.22	
Enterprise capital income	0.10 (0.18)	0.29	
Enterprise taxes paid	0.04 (0.03)	0.07*	
Panel C: Expenditure and income multipliers			
Average of both multipliers	2.60 (1.44)	0.04**	0.12
Joint test of both multipliers		0.01**	0.06*

Notes: This table is analogous to Table 5 (see table notes for detail). The only difference is that here, monetary values are nominal, whereas in Table 5 transfer amounts and outcome variables are deflated to January 2015 using the overall consumer price index in the geographically closest market. Standard errors and test statistics are computed from 2,000 replications of a wild clustered bootstrap. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

I Study pre-analysis plans

We filed a series of pre-analysis plans as part of this study. These include: i) Haushofer et al. (2017a), and a follow-up amendment outlining spillover analyses, both of which focus on household outcomes; ii) Haushofer et al. (2016), which covered midline market price and enterprise data; and iii) Haushofer et al. (2018), which focused on macroeconomic quantities of interest. All pre-analysis plans can be accessed on the AEA trial registry: `https://www.socialscienceregistry.org/trials/505`. In this paper, we focus on primary outcomes for households, enterprises and prices, collected as part of our baseline and endline household and enterprise censuses and surveys, as well as our midline market price surveys.

Less relevant to this paper are: i) Walker (2017), which forms the basis of Walker (2018) on local taxes and public goods; and ii) Haushofer et al. (2017b), which conducts a separate exercise to study potential transfer targeting.

In the interest of space, we do not present an exhaustive list of every outcome component and analysis mentioned across these pre-analysis plans. A supplemental appendix containing the full set of pre-specified outcomes for these plans is available online at `https://osf.io/r5q6v/`.

Table I.1 presents the 10 primary household outcomes that we pre-specified as part of a single table, including FDR q-values accounting for multiple testing across these ten outcomes. In addition to the specifications reported in the main tables, we also report the pooled saturation effect, the average effect of being in a high saturation sublocation across all eligibility and village types. As outlined in Section 3, we prefer our spatial estimates as they take advantage of the full variation in treatment intensity in our data, but present these saturation results for completeness.

To calculate the pooled saturation effect, we use coefficient estimates from the following equation:

$$\begin{aligned} y_{hvs,t=1} = \beta_0 + \beta_1 T_{vs} + \beta_2 E_{hvs} + \beta_3 H_s + \beta_4 T_{vs} \times E_{hvs} + \beta_5 T_{vs} \times H_s \\ + \beta_6 E_{hvs} \times H_s + \beta_7 T_{vs} \times E_{hvs} \times H_s + \delta_1 y_{hvs,t=0} + \delta_2 M_{hvs} + \varepsilon_{ihvs}. \end{aligned} \tag{14}$$

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. 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; $\times$ denote interaction terms. Standard errors are clustered at the saturation group level. The pooled saturation effect is then a weighted average of $\beta_3, \beta_5, \beta_6$ and β_7 using population weights of all households across high-saturation sublocations.

We make two additional notes. First, in Haushofer et al. (2016), we were not clear whether we would focus on a balanced panel of market survey data or an unbalanced panel. For simplicity, we present results using a unbalanced panel, but results are robust to using a balanced panel. Second, our reduced form equations cluster standard errors at the village level, as pre-specified in Haushofer et al. (2017a), but results are also robust to clustering at the sublocation level. (Both sets of results are available upon request).

Table I.1: Pre-specified primary outcomes, household welfare plan

	(1)	(2)	(3)	(4)	(5)
	Recipient Households		**Non-Recipient Households**		
	𝟙(Treat village) Reduced form	Total Effect IV	Total Effect IV	Pooled saturation effect	Control, low saturation mean (SD)
Assets (non-land, non-house), net borrowing	178.09***	182.01***	132.63*	41.96	1,132.15
	(24.61)	(44.25)	(78.32)	(42.33)	(1,420.22)
	[0.00]***	[0.00]***	[0.22]	[0.48]	
Household expenditure, annualized	292.98***	338.16***	333.73***	138.26*	2,536.86
	(60.09)	(109.36)	(123.24)	(71.29)	(1,934.09)
	[0.00]***	[0.01]***	[0.05]*	[0.27]	
Household income, annualized	77.62*	134.02	229.46***	111.85*	1,023.45
	(43.66)	(93.83)	(88.59)	(59.47)	(1,634.70)
	[0.07]*	[0.16]	[0.05]*	[0.27]	
Household revenue, annualized	73.62	175.46*	54.32	117.04**	933.27
	(51.59)	(90.98)	(109.21)	(57.86)	(1,698.65)
	[0.12]	[0.13]	[0.38]	[0.27]	
Psychological well-being index	0.09***	0.12*	0.08	0.04	0.01
	(0.03)	(0.07)	(0.06)	(0.03)	(1.01)
	[0.01]**	[0.13]	[0.27]	[0.27]	
Health index	0.04	0.06	0.01	−0.01	0.03
	(0.03)	(0.06)	(0.05)	(0.03)	(1.01)
	[0.14]	[0.26]	[0.41]	[1.00]	
Education index	0.09**	0.09*	0.10*	0.03	0.01
	(0.04)	(0.05)	(0.06)	(0.03)	(1.02)
	[0.02]**	[0.13]	[0.22]	[0.50]	
Female empowerment index	−0.01	0.08	0.09	0.02	0.05
	(0.07)	(0.14)	(0.15)	(0.08)	(0.94)
	[0.35]	[0.34]	[0.37]	[1.00]	
Food security index	0.10***	0.05	0.08	0.01	0.01
	(0.03)	(0.07)	(0.06)	(0.03)	(1.00)
	[0.01]***	[0.34]	[0.25]	[1.00]	
Hours worked last week (respondent)	1.27	−1.87	−1.79	1.05	34.05
	(1.02)	(1.85)	(1.41)	(0.95)	(27.11)
	[0.14]	[0.26]	[0.29]	[0.48]	

Notes: Each row represents regressions of a pre-specified primary outcome on different regressors. Column 1 reports the coefficient on an indicator for treatment village from a regression using data from eligible households (as classified by the GE census team), and includes an indicator for saturation status of the sublocation (Equation 1). Column 2 reports the total effect on treated households (eligible recipients) from the "optimal" IV spatial regression of each outcome on the amount transferred per capita to a household's own village v (instrumented by village treatment status), and to villages other than v in each 2km radii band around the household (instrumented by the share of eligible households assigned to treatment in villages other than v inside the buffer), as in Equation 2. For this analysis, the sample is restricted to eligible households. We have 5,168 to 5,423 observations (1,118 for the female empowerment index) for these columns. Column 3 presents the average spillover effect on eligible households in control villages as well as ineligible households, coming from a stacked spatial regression of each outcome on the amount transferred per capita GDP to each 2km radii band around each household (instrumented by the share of eligibles assigned to treatment in each buffer), as in Equation 3. The reported average effect comes from a population-share-weighted average effect experienced by those two groups, and is representative of the average untreated household. We have 5,230 to 5,509 (978 for female empowerment) for column 3.. The number of radii bands included in columns 2 and 3 is chosen, as pre-specified, by minimizing the BIC. Column 4 reports the pooled saturation effect, the average saturation effect experienced by households in high-saturation sublocations, derived as a weighted average of $\beta_3, \beta_5, \beta_6$ and β_7 in Equation 14. We have to 7,832 to 8,239 (1,535 for female empowerment). Column 5 reports the weighted mean and standard deviations of the outcome variables in low-saturation control villages (across eligible and ineligible households). Each regression is weighted by inverse sampling weights and contains baseline values of the outcome when available. Standard errors are clustered at the village in column 1, at the sublocation level in column 4, and calculated following Conley (2008) using a uniform kernel out to 10 km in columns 2 and 3. Minimum FDR q-values are reported in brackets. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

J Additional welfare analysis

We provide here an illustrative mapping between household welfare and aggregate output, emphasizing that any increase in aggregate output must reflect some combination of (i) an increase in the *employment* of factors of production, which comes at an opportunity cost, and (ii) an increase in their aggregate *productivity*, which does not. We also illustrate how household welfare differs from household expenditure.

Consider a household i whose market interactions at time t involve the (net) purchase of a vector of commodities c_{it} at prices p_t, the supply of a vector l_{it} of labor services at wages w_t, and the supply of (net) savings to support capital investment equal to the difference between current-period income and expenditure. In addition, the household receives profit π_{it} from owned enterprises.[52] The household's problem is

$$\max_{\{c_{it},l_{it}\}} u(\{c_{it},l_{it}\},\{c_{-it},l_{-it}\}) + \lambda_i \left(T_i + \sum_{t=0}^{\infty} \delta^t (\pi_{it} + w_t \cdot l_{it} - p_t \cdot c_{it}) \right) \tag{15}$$

where λ_i is the Lagrange multiplier on the budget constraint and $\delta \equiv 1/r$ the discount rate on future funds.[53] The economy's capital stock at the beginning of period t is $k_t = \sum_i k_{it} = \sum_i \sum_{\tau=-\infty}^{t-1} \delta^{\tau}(\pi_{i\tau} + w_t \cdot l_{i\tau} - p_t \cdot c_{i\tau})$. The household's contribution to (real) output in period t measured using the income approach is equal its claims on firm profits plus the factor payments it receives, or $\pi_{it} + w_t \cdot l_{it} + r \cdot k_{it}$.[54]

Overall output is the sum of these contributions which is simply total enterprise value added, and the period-t contribution to the transfer multiplier is the effect of $1 of transfers on this quantity. Whether distributed to households in the form of higher profits, wages, or interest payments, real output gains can be achieved only through increases in (i) the supply of labor or capital, or (ii) of productivity. In the case of labor supply this comes at a utility cost, since $\frac{\partial u}{\partial l_{it}} < 0$, so that a dollar increase in output must be worth less than a dollar in equivalent variation terms. Similarly in the case of capital, an increase in the period-t capital stock implies a decrease in consumption in some other period(s), so that again a dollar increase in output is worth less than a dollar in equivalent variation terms. In the case of a pure productivity gain, on the other hand, a dollar of output is worth a full dollar to the household(s) that receive it.

To contrast expenditure with welfare, assume for simplicity that first-order conditions are necessary and sufficient for a solution to the household's problem defined by (15), and

52. The term π_{it} could also capture other (net) transfers e.g. from peer households and from the government. We ignore these terms here as the estimated treatment effects on them in our data are negligible.

53. One can generalize this formulation to allow for non-separable household production using non-marketed inputs such as family labor without changing the basic message. It is also straightforward to allow for discount rates to vary across agents, reflecting capital market imperfections.

54. Its contribution measured using the consumption approach is its expenditure $p_t \cdot c_{it}$ plus its attributable share of firm investment which (assuming a closed economy) must equal household savings, i.e., $p_t \cdot c_{it} + (\pi_{it} + w_t \cdot l_{it} + r \cdot k_{it} - p_t \cdot c_{it})$, which is evidently equivalent.

for simplicity we ignore externalities. The envelope theorem then implies

$$\frac{dv_i}{dT} = \lambda_i \sum_{t=0}^{\infty} \delta^t \left(\frac{t}{\delta} \frac{\partial \delta}{\partial T} s_{it} + \frac{\partial \pi_{it}}{\partial T} + \frac{\partial w_t}{\partial T} \cdot l_{it} - \frac{\partial p_t}{\partial T} \cdot c_{it} \right) \tag{16}$$

where $s_{it} = \pi_{it} + w_t \cdot l_{it} - p_t \cdot c_{it}$ is period t savings. In comparison, the welfare effect of a marginal change in T_i holding other transfers fixed is simply

$$\frac{\partial v_i}{\partial T_i} = \lambda_i \tag{17}$$

The (marginal) equivalent variation dEV_i/dT is the ratio of these expressions. To see how this relates to household expenditure, define (the present discounted value of) household expenditure as $e_i = \sum (1/(1+r))^t p_t \cdot c_{it}$. Differentiating the budget constraint, we have

$$\frac{de_i}{dT} = \sum_{t=0}^{\infty} \delta^t \left(\frac{t}{\delta} \frac{\partial \delta}{\partial T} (\pi_{it} + w_t \cdot l_{it}) + \frac{\partial \pi_{it}}{\partial T} + \frac{\partial w_t}{\partial T} \cdot l_{it} + w_t \cdot \frac{\partial l_{it}}{\partial T} \right) \tag{18}$$

Comparing the equations above, we see that

$$\frac{dEV_i}{dT} = \frac{de_i}{dT} - \sum_{t=0}^{\infty} \delta^t \left(\frac{t}{\delta} \frac{\partial \delta}{\partial T} (p_t \cdot c_{it}) + \frac{\partial p_t}{\partial T} \cdot c_{it} + w_t \cdot \frac{\partial l_{it}}{\partial T} \right) \tag{19}$$

This expression shows that changes in expenditure are closely related to changes in equivalent variation, but with several intuitive (and correctable) sources of bias. First, (nominal) expenditure incorrectly counts appreciation of the price of the household's planned time path of consumption, whether due to appreciation of intra-period prices ($\frac{\partial p_t}{dT} \cdot c_{it}$) or of the inter-period interest rate $\frac{t}{\delta} \frac{\partial \delta}{\partial T} (p_t \cdot c_{it})$, as a welfare gain. This is why constant-dollar expenditure measures are preferable. Second, it incorrectly counts income gains due to behavioral responses such as increased labor supply ($w_t \cdot \frac{\partial l_{it}}{\partial T}$) as a welfare gain. Finally, if (more realistically) we were to examine expenditure over any finite period of time this would introduce a third bias, as this metric would count as a welfare gain any increases in current expenditure that were driven by decreases in future expenditure (i.e., by dis-saving).

References

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate." *Biometrika:* 491–507.

Conley, Timothy G. 2008. "Spatial Econometrics." In *The New Palgrave Dictionary of Economics,* Second Edition, edited by Steven N. Durlauf and Lawrence E. Blume, 7:741–47. Houndsmills: Palgrave Macmillan.

de Mel, S., D.J. McKenzie, and C. Woodruff. 2009. "Measuring microenterprise profits: Must we ask how the sausage is made?" *Journal of Development Econ.* 88 (1): 19–31.

Deaton, Angus, and Salman Zaidi. 2002. *Guidelines for constructing consumption aggregates for welfare analysis.* Vol. 135. World Bank Publications.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *QJE* 131 (2): 799–858.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker. 2020. "Pre-analysis Plan Report for General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya." Available at .

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

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

———. 2017b. "GE Effects of Cash Transfers: Pre-analysis plan for targeting analysis." September.

———. 2018. "General Equilibrium Effects of Cash Transfers: Pre-analysis plan." June.

Haushofer, Johannes, and Jeremy Shapiro. in prep. "The Short-term Consumption Response to Unconditional Cash Transfers."

———. 2013. "Policy Brief: Impacts of Unconditional Cash Transfers." Policy Brief, October.

———. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *QJE* 131 (4): 1973–2042.

Walker, Michael. 2017. "Pre-Analysis Plan: Local Public Finance and Unconditional Cash Transfers in Kenya." February.

———. 2018. "Informal Taxation Responses to Cash Transfers: Experimental Evidence from Kenya." July.