Cash versus Kind:

Benchmarking a Child Nutrition Program against Unconditional Cash Transfers in Rwanda*

Craig McIntosh[†] and Andrew Zeitlin[‡]

Revised; November 30, 2023

Abstract

We develop a methodology to benchmark in-kind programs against cost-equivalent cash transfers. Our application compares a multi-dimensional child nutrition intervention to unconditional cash transfers, using randomized variation in transfer amounts and regression adjustment of expenditures to estimate impacts of cash transfers at identical cost as well as to estimate the return to increasing cash transfer amounts. While neither the in-kind program nor a cost-equivalent transfer costing $124 per household moves core child outcomes within a year, cash transfers create significantly greater consumption and asset accumulation. A larger cash transfer costing $517 substantially improves consumption and investment outcomes and drives modest improvements in dietary diversity and child growth.

Keywords: Experimental Design, Cash Transfers, Malnutrition

JEL Codes: O12, C93, I15

*We are grateful to DIV, Google.org, and USAID Rwanda for funding, and to USAID, CRS, GiveDirectly, CEGA/DIL, and IPA for their close collaboration. We thank Leodimir Mfura and Marius Chabi for overseeing the fieldwork, and Richard Appell, Sarait Cardenas-Rodriguez, Chris Gray, Ali Hamza, and Bastien Koch for research assistance. This project is covered by Rwanda National Ethics Committee IRB 143/RNEC/2017 and IPA IRB 13730, and the study is pre-registered with the AEA as trial AEARCTR-0002559. This study is made possible by the support of the American People through the United States Agency for International Development (USAID). The contents of this study are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

[†]University of California, San Diegoctmcintosh@ucsd.edu

[‡]Georgetown University, andrew.zeitlin@georgetown.edu

1 Introduction

Should governments and aid agencies provide assistance in kind, or would beneficiaries be better off if they just received the money? Unconditional cash transfers provide a natural benchmark to in-kind programs for several reasons. From a practical perspective, cash transfers generate well-documented benefits for households (Aizer et al., 2016; Haushofer and Shapiro, 2016), and their costs are falling across the developing world due to the penetration of mobile money (Suri, 2017). Cash programs are flexible and can be rescaled in terms of targeting or expense, which allows them to be paired at cost-equivalent levels against more complex interventions.[1] By contrast, comparisons across programs in different contexts and at different costs rely on strong assumptions about external validity (Vivalt, 2015) and linearity (both in cost-effectiveness ratios and the preferences of policymakers across individuals). More fundamentally, decisions made over the use of unconditional transfers may reveal information about the preferences of the beneficiary households that is important in interpreting the welfare effects of more complex programs (Finkelstein and Hendren, 2020). While missing markets, externalities, or other failures can justify the provision of in-kind programs, benchmarking the impact of such programs against cost-equivalent cash transfers forces policymakers to be explicit about the circumstances that would merit a more multidimensional, overhead-heavy approach. For these reasons, recent years have seen increasing calls for head-to-head studies that use cash as a benchmark for more complex forms of international development assistance.[2]

In this study we develop an experimental approach to *cost-equivalent benchmarking* of in-kind interventions to cash transfers, with an application to a nutritional and maternal health intervention in Rwanda. The in-kind program, Catholic Relief Service's *Gikuriro*, tackles malnutrition from multiple dimensions, resulting in a highly bundled intervention. Since cash transfers are have been shown in other contexts to drive child consumption outcomes (Aguero et al., 2006; Baird et al., 2019; Seidenfeld et al., 2014), this context and set of outcomes represent an ideal setting for the

[1] In medical research, new interventions are typically benchmarked against the best current treatment, referred to as the 'standard of care'. Perhaps the best-validated anti-poverty tool in the development literature is BRAC's Targeting the Ultra-Poor program (Banerjee et al., 2015); however given that this program is expensive ($3,700 per household), lengthy in duration (a year of intensive intervention), and targeted at very specific households (the ultra-poor), it represents a somewhat inflexible benchmark for the broad range of potential development interventions.

[2] In particular Blattman and Niehaus (2014) argue for the use of cash transfers as the 'index funds' of international development, providing a reference rate of return that could be used to hold donors accountable.

benchmarking comparison. *Gikuriro's* multi-faceted approach has been broadly advocated in the public health literature (Ruel et al., 2013) and has been shown to be effective in similar contexts (Leroy et al., 2016).[3] Such programs attract substantial funding worldwide; for example, USAID's Bureau for Humanitarian Assistance invests approximately 400 million dollars annually to support similar multi-sectoral food security programs.[4] This necessarily complex and overhead-intensive approach presents a stark contrast with the pared-down nature of a cash-transfer alternative in which the share of spending received by the beneficiaries is maximized, implemented here by the U.S. non-profit GiveDirectly. Further, our study districts feature under-five stunting rates of 35–45 percent (DHS, 2016), portending dire long-term consequences for human capital formation (Currie and Almond, 2011). This context therefore provides two contrasting approaches with distinct causal pathways to the same, critical policy outcome.

To form this comparison, we conduct a cluster-randomized trial across 248 villages located in Kayonza and Nyabihu districts of Rwanda.[5] We enroll households into the study if they are defined as poor by the Government of Rwanda, and contain either pregnant women or children under five years of age. Eligible beneficiaries were assigned at the village level to be offered either the Gikuriro intervention or a cash transfer. Given that the eventual exact costs of programs are not known *ex ante*, we randomized cash transfer values around the expected cost; this variation is then combined with an *ex post* linear regression adjustment of the costs across arms to form an exact, cost-equivalent comparison. Thirteen months after baseline, we measure impacts on five primary outcomes—household consumption, household dietary diversity, child and maternal anemia, child growth, and household non-land net wealth—as well as a broader set of secondary measures. These outcomes are documented as core drivers of improved long-run outcomes for children (Hoddinott et al., 2013; Maluccio et al., 2009), as well as being plausibly impacted by both intervention types.[6]

[3]The *Lancet*'s Maternal and Child Nutrition Study Group specifically advocates for this type of multidimensional approach, saying "[t]he need for investments to boost agricultural production, keep prices low, and increase incomes is indisputable; targeted agricultural programmes can complement these investments by supporting livelihoods, enhancing access to diverse diets in poor populations, and fostering women's empowerment" (Ruel et al., 2013).

[4]Most of these programs include infant and child feeding practices, maternal health and nutrition within the 1000 days window, WASH, village saving and loan associations, livelihoods and agriculture, and disaster risk management components. Currently, they have similar programs in Bangladesh, Niger, Burkina Faso, Mali, DRC, Uganda, Malawi, Ethiopia, Kenya, Madagascar, Zimbabwe, and Haiti.

[5]The village-level study design was motivated by the clustered nature of the Gikuriro intervention. In addition, this helps to allay concerns about the potential for spillovers of cash transfers.

[6]All primary and secondary outcomes were registered prior to receipt of endline data on the American Economic Association RCT Registry, with ID AEARCTR-0002559.

The findings of this study design, like that of a companion study that benchmarks cash against a youth employment program (McIntosh and Zeitlin, 2022), allow us to assess tradeoffs across *outcomes*, across *sub-populations*, and across *expenditure per beneficiary*.[7] Across outcomes, the cash and in-kind programs vary in their relative efficacy. At a cost of $142, Gikuriro was successful at delivering gains in savings, a domain that was the target of one core program component. It did not lead to improvements in consumption, dietary diversity, wealth, child anthropometrics, or anemia within the thirteen-month period of the study. A cost-equivalent cash transfer has significantly larger effects on consumption, allows households to pay down debt, and generates increased investment in productive and consumption assets. Across sub-populations, we uncover very little evidence of heterogeneity in impacts, suggesting that targeting these interventions more tightly would not increase the average effect substantially. Across expenditure levels, our results show that a much larger cash transfer costing $567 led to across-the-board improvements in consumption-based welfare measures, a substantial improvement in dietary diversity, a drop in child mortality, and modest improvements of about 0.1 standard deviation in the anthropometric indicators of height-for-age, weight-for-age, and mid-upper arm circumference (all significant at 10 percent or above). The large cash transfer delivers benefits even on outcomes specifically targeted by the other program. While it is unsurprising that very large amounts of money show up in consumption and productive assets, the improvements in diet and particularly child anthropometrics over such a short period of time are impressive (although our study does not feature an arm with a high-cost implementation of the in-kind arm to correspond to the large cash transfer). For a given target population, policymakers must therefore decide how to trade off between relatively intensive interventions that can only be provided to a small subset, or lower-cost interventions that can reach a larger proportion of this target group.[8]

The primary contribution of this paper is to develop a methodology that allows for exact cost-equivalent comparisons in the face of the inevitable *a priori* uncertainty that exists as to program costs. This is critical because comparison of programs operated at different costs requires an as-

[7]The successor study uses the same benchmarking methodology (comparing an in-kind program to randomized unconditional cash transfers) to answer an applied question from the labor literature: namely, how best to help underemployed youth gain access to productive work.

[8]A fourth dimension of tradeoff, namely across time scales, is not possible in this paper because the control group was treated immediately after the one-year endline. In the Conclusion we discuss the implications of the relatively short duration of this study as well as the asymmetry arising from the fact that cash transfer flows begin shortly after baseline, while the in-kind arm was delivered in the middle of this transfer period.

sumption of linearity-in-expenditure that is not typically testable within each arm. Alternatively, studies that attempt to generate cost equivalent comparisons based on anticipated costs at the time of design will almost inevitably be wrong after the fact, undermining their core purpose. By randomizing cash transfer amounts we can achieve two distinct goals, both allowing for cost adjustment based on the correct, *ex post* costs, and examination of the way that the returns to cash vary with spending per person (that is, the question of linearity-in-expenditure). Our approach yields an ex-post cost-equivalent comparison with relatively little power loss relative to a design that takes a single ex-ante guess about these ex-post costs, while permitting robustness checks that test sensitivity to the nature of the interpolation.

In applying this approach, we extend the small number of studies using cash transfers as one arm within multi-armed trials. The largest extant literature is based on the comparison of cash aid to food aid (Ahmed et al., 2016; Hidrobo et al., 2014; Hoddinott et al., 2014; Leroy et al., 2010; Schwab et al., 2013), uncovering a fairly consistent result that food aid leads to a larger change in total calories while cash aid leads to an improvement in the diversity of foods consumed, while benefits for targeted individuals can be limited and there can be adverse consequences for non-targeted individuals when market imperfections mean that cash transfers cause food price increases (Cunha et al., 2019; Filmer et al., 2021).[9,10] Efforts to compare more complex, multi-dimensional programs against cash include BRAC's Targeting the Ultra-Poor program (Chowdhury et al., 2017), microfranchising (Brudevold-Newman et al., 2017), agricultural inputs (Brudevold-Newman et al., 2017), and sustainable livelihoods (Sedlmayr et al., 2017). These studies have typically struggled with the question of how to anticipate costs and compliance well enough to realize an exact cost-equivalent comparison after the fact; we address this with our study design. Further, we extend the existing literature comparing in-kind interventions to cash by illuminating the extent of heterogeneity in the returns to cash versus kind, and by showing how the comparison of in-kind programming with a larger cash transfer on a benefit-cost basis would lead to different conclusion.

[9] An interesting contribution to this food versus cash literature is Hirvonen and Hoddinott (2020), who find that poor and geographically marginalized households may prefer food transfers because of a lack of consumer markets during times of shortfall.

[10] A related literature compares food transfers with *vouchers* redeemable for the purchase of food. Recent work by Banerjee et al. (2021) shows positive effects of shifting to vouchers, attributable to changes in the share of amount of assistance reaching targeted households and in the quality of food consumed.

Beyond this, we add to the growing number of studies that show meaningful impacts of cash transfers on important life outcomes in the short term, in domains with plausible channels for long-term impact. These include evidence of impacts on child nutrition (Aguero et al., 2006; Seidenfeld et al., 2014), schooling (Skoufias et al., 2001), mental health (Baird et al., 2013; Samuels and Stavropoulou, 2016), teen pregnancy and HIV (Baird et al., 2011), microenterprise outcomes (De Mel et al., 2012), consumer durables (Haushofer and Shapiro, 2016), and productive assets (Gertler et al., 2012). The evidence on the long-term impacts of cash transfers is more mixed, but some studies have found substantial impacts (Aizer et al., 2016; Barham et al., 2014; Fernald et al., 2009; Hoynes et al., 2016).[11] We contribute to the cash transfer literature by evaluating multiple transfer amounts and modalities in the same context, and illuminating the scope for beneficiary choice over cash-transfer modality to drive impacts.

The rest of the paper is organized as follows. Section 2 lays out the study design. In Section 3, we present simple ITT analyses of the experimental results. presents the core empirical results of the benchmarking exercise, In Section 4, we formalize the policy choice problem, undertake our cost-equivalent approach to benchmarking policy comparisons, and compare this to results of a traditional cost-effectiveness approach to comparative evaluation. Section 5 concludes.

2 Study design

2.1 Interventions studied

The Gikuriro program was developed by USAID, Catholic Relief Services (CRS), and the Netherlands Development Organization (SNV) to "improve the nutritional status of women of reproductive age and children under five years of age, with an emphasis on the 1,000 day window of opportunity from pregnancy until a child's second birthday" (Figlio et al., 2014).[12] The resulting multi-faceted program brings together several components in order to attack this problem from multiple directions at once, and is a central pillar of the Government of Rwanda's approach to combatting malnutri-

[11] For examples of studies that find dissipating long-term benefits, see Baird et al. (2016), Araujo et al. (2017), and Blattman et al. (2020). For evidence from systematic reviews of cash transfers on schooling see (Molina-Millan et al., 2016), and on child health see (Manley et al., 2013; Pega et al., 2014).

[12] Gikuriro means "well-growing child" in Kinyarwanda. Its objectives are documented in USAID Cooperative Agreement No.: AID-696-A-16-00001, 2015).

tion in rural Rwanda.[13] Gikuriro combines an integrated nutrition program with a standard WASH curriculum (water, sanitation, and hygiene), and seeks to build the capacity of the health infrastructure providing services to mothers and newborns, particularly Community Health Workers (CHWs). The program also seeks to improve livelihoods by providing additional assistance to eligible households, including (a) Village Nutrition Schools (VNS); (b) Farmer field learning schools (FFLS), which potentially includes distribution of small livestock, fortified seed, etc.; (c) Savings and Internal Lending Communities (SILCs); and (d) the Government of Rwanda's Community-Based Environmental Health Promotion Program (CBEHPP). In addition, Gikuriro provided a program of Behavioral Change Communication (BCC), supporting participation in all components of the program including savings, agriculture, and nutrition, as well as hygiene. Similarly designed programs in neighboring countries have been shown to have positive impacts.[14] This comprehensive approach seeks to build supply and demand for child health services simultaneously, and is fairly typical of the kinds of multi-sectoral child health programs implemented by USAID in many parts of the developing world.[15]

In terms of program expenditures, the most substantial component of Gikuriro is the Farmer Field Schools, which consume 41% of the overall budget. The main cost driver here was the distribution of seeds, as well as small livestock and poultry. Next most important were the village nutritional schools, with 19% of the spend. This component's core goal was to use demonstration plots to show how to use very small plots (or even gunny sacks for the land poor) to grow micro-gardens with a variety of nutrient rich greens that could be used to support child nutrition. The other components of program expenditure were overall logistical program support (22%), monitoring and evaluation (9%), the SILCs (6%), and the BCC and WASH trainings (4%). Our estimates

[13] USAID's Global Health and Nutrition Strategy explicitly calls for multi-sectoral interventions that incorporate agriculture, WASH, education, and outreach to mothers in the first 1,000 days through the public health system. The agency reports reaching 27 million children worldwide under the age of 5 in 2016 alone through such programs, which represent the prescribed USAID modality for Scaling up Nutrition (SUN) countries.

[14] Leroy et al. (2016) estimate the impacts of *Tubaramure*, a program in Burundi undertaken by the same implementer, which provided food supplementation along with a BCC strategy akin to that of Gikuriro, finding reductions in anemia among post-partum women and improvements in dietary diversity. Perhaps closest in design to Gikuriro is the combined WASH and nutrition intervention studied in rural Kenya by Null et al. (2018). Like Gikuriro, they studied the impacts of a combined WASH and nutrition intervention. They found that the nutrition intervention (whether offered together with WASH or alone) delivered improvements of between 0.11 and 0.14 Z scores of weight-for-age, with "almost all of the growth benefits in the nutrition group and combined water, sanitation, handwashing, and nutrition group were already manifest in the first year."

[15] Examples of similar integrated WASH/agriculture/child nutrition programs funded by USAID include SPRING in Bangladesh, RING in Ghana, Yaajende in Senegal, and ENGINE in Ethiopia.

suggest that a household participating in every component of Gikuriro would have actually received training and inputs worth \$70.13, of which \$5.06 is direct transfer of materials and inputs.

To benchmark the impact of this program against cash we worked with GiveDirectly, a US-based 501(c)3 Non-Profit organization. GiveDirectly specializes in sending mobile money transfers directly to the mobile phones of beneficiary households to provide large-scale household grants in developing countries including Kenya, Uganda, and Rwanda. GiveDirectly's typical model has involved targeting households using mass-scale proxy targeting criteria such as roof quality. GiveDirectly builds an in-country infrastructure that allows them to enroll and make transfers to households while simultaneously validating (via calls from a phone bank) that transfers have been received by the correct people and in a timely manner. Their typical transfers are large and lump-sum, on the order of \$1,000, and the organization provides a programatically relevant counterfactual to standard development aid programs because it has a scalable business model that would in fact be capable of providing transfers to the tens of thousands of households that are served by the Gikuriro program. Because of the nutritional focus of the Gikuriro intervention, GiveDirectly incorporated a 'nudge' into the way the program was introduced (Benhassine et al., 2015), distributing a flyer at the time of the intervention that emphasized the importance of child nutrition. An English translation of this flyer is included in Online Appendix A.2. Given observed impacts of cash transfers on other goods, e.g., productive assets and housing value, it is evident that households felt at liberty to spend the grants on items not directly related to child nutrition.

2.2 Study outcomes and eligibility

We pre-committed to five primary outcomes and seven secondary outcomes for the study. The primary outcomes are consumption, dietary diversity, child growth (height-for-age and weight-for-age, as well as mid-upper arm circumference), and household non-land wealth. The secondary outcomes are borrowing and saving, pregnancy and live birth rates, health knowledge, mortality, health-seeking behaviors, productive assets, and housing quality. A more detailed description of all outcomes is provided in the Online Appendix. Data were collected by Innovations for Poverty Action (IPA) teams through two survey waves; the baseline was conducted in August and September of 2016, and the endline was conducted in September and October of 2017. The baseline included a comprehensive household survey module as well as anthropometric measurement of all children five and under in eligible households.[16] The endline consisted of these same measures plus blood-based anemia testing for the study children and for all women of childbearing age.[17]

Both implementers made contact with the study subjects and began enrollment immediately after baseline. Give Directly (GD) began implementation shortly after the baseline meaning that at endline individuals in that arm had experienced about 12 months of the household grants treatment (running up through the month before endline). Gikuriro was slower than the cash program to begin implementation on the ground; in that arm households had typically experienced 8-9 months of household-level implementation at the time of the endline.[18] The duration of the RCT component of the study was limited by the fact that local governments wanted to hit targets for the broader, national rollout of nutritional and WASH programming, and hence we were not able to maintain the control groups for more than one year. We cannot therefore speak to the long-term impacts of the interventions. Anticipating this issue, we took two approaches to measurement. One of them was to try capture the stocks of intertemporal assets that would be the obvious conduits to future consumption benefits for the households. The second was to emphasize outcomes such as dietary diversity and anemia that have the potential to respond quickly to changes in consumption patterns, while also retaining the more standard metrics of child malnutrition such as height for age (HAZ), weight for age (WAZ), and mid-upper arm circumference (MUAC).[19] Further, a number of recent RCTs have shown that programs can have meaningful impacts on biometric outcomes over timeframes similar to that analyzed in this study, such as Desai et al. (2015), Leroy et al. (2016), Fink et al. (2017), and Null et al. (2018).

A major practical issue with this kind of comparative study is how to define the the target group, given that each implementer would naturally do this differently. We collectively resolved this by agreeing to a set of readily observable characteristics that both implementers deemed eligible for their funding and desirable to target, and then tasking the survey firm with listing all households

[16] We weighed all children younger than 6 years once using a Seca 385 scale. We measured length for children under two years with Seca 417 length boards and height for older children to the nearest 0.1 cm with Seca 213 stadiometers. These were measured twice and we took the average of the two measurements unless they differed by more than 0.7 cm in which case we took a third measurement and averaged the two closest measurements.

[17] Using the guideline for anemia testing in population-based surveys, we used HemoCue Hb 301 system.

[18] Since both programs had six months of notice that they would be implementing in the study sample in these two districts and began national-level implementation at the same time, this differential delay likely reflects a real difference in the relative ramp-up speeds of cash versus more complex programming.

[19] Dietary diversity is an immediate indicator of improvements in consumption, and the clinical literature has shown that anemia tests respond within 3 months of improvements in diet (Habicht and Pelletier, 1990).

in study villages and allocating them to the 'eligible' or 'ineligible' stratum. Households were considered eligible if they contained a child known to the government growth monitoring system to be malnourished, or else were in the poorest two government poverty classifications (Ubudehe 1 or 2) and contained either children under the age of 5 or a pregnant/lactating woman. All other households in study villages are classified as ineligible. The survey firm passed the names of all eligible households to the implementing partners in an identical way, and we then randomly sampled up to 8 eligible and 4 ineligible households per village for the study. The identity of the sampled households was not revealed to either implementer. In this paper we use survey weights to estimate impacts that are representative for all eligible households in study villages.

Both implementers concurred quite closely with our definition of eligibility on the ground, and compliance was high: we have 80 percent of the eligibles treated by Gikuriro and 84 percent by GiveDirectly.[20]

2.3 Assignment Protocol

Randomization occurred at the village level across 248 villages, using a blocked randomization where the blocks were formed by the combination of districts and village-level poverty scores within district, creating a total of 22 blocks with between 10 and 13 villages per block. Fixed effects for these blocks are included in all analysis. A computer was used to conduct the randomization based on a frame of villages agreed to by CRS and government officials.

Table 1 presents a schematic of the design of the study. 74 villages were assigned to the Gikuriro intervention, 74 were assigned to the control group (no intervention), and 100 were assigned to GiveDirectly household grants. The GiveDirectly villages were further split into four transfer amounts, randomized at the village level. Three treatment amount arms, with 22 villages in each, received transfer amounts in a range around the anticipated cost of Gikuriro. A final 34 villages were assigned to the 'large' GiveDirectly transfer amount which was selected by GiveDirectly as the amount anticipated to maximize the cost effectiveness of cash. At the individual level, those in the cash arm were assigned to receive their transfers either as a "flow" of monthly transfers over

[20] Because eligibility status was determined from records, in some cases the Ubudehe status of the households proved to be incorrect or unverifiable when they were approached by GD for treatment, and hence were not offered the program. Gikuriro implemented a consultative process with community members that formed the basis of their targeting.

twelve months (75 percent of participants); as a lump-sum transfer in the first month (17.5 percent of participants); or as a choice between these. The transfers actually received by households in the GD 'main' arms were $41.32, $83.63, and $116.91; the large GD arm actually transferred $532 to households. All transfer amounts were translated into Rwandan Francs at an exchange rate of 790 RwF per US dollar, and were rounded to the nearest hundred.[21] Appendix Figure A.1 provides a box and whisker plot of the randomly assigned mean transfer amount per village relative to the actual amount received per household observed in the GD institutional data, and shows that the two correspond closely.

2.4 Compliance

Compliance with assigned treatments is incomplete. This has implications for both the assignment of costs per beneficiary and the interpretation of impacts, as discussed in Section 2.5 below.

While defining compliance with cash transfers is conceptually straightforward, compliance with Gikuriro is complex because of the multi-dimensional nature of the program. We have five forms of participation that can be ascribed directly to the program: participation in three types of training (nutritional, cooking/hygiene, and agricultural extension), whether households have themselves harvested the nutrient-rich household vegetable plots that they were trained by the Farmer Field Schools to grow, and whether they received livestock directly from Gikuriro.

Appendix Table A.1 examines the determinants of participation with specific sub-components of Gikuriro, based on self-reports within the eligible population, as well as with cash transfers as recorded by GiveDirectly. Among the representative sample of eligible households, 63% received nutritional training, 51% cooking training, 57% farmer training, 48% grew home vegetable gardens, and 33% received livestock. On average eligible households engaged in 2.5 of these activities, and well over three quarters of households engaged in at least one of them. Even among the entire village population, 40% participated in some part of Gikuriro. The two driving determinants that

[21] GiveDirectly believed that the most cost-effective use of these funds would be to attempt to equalize the amount transferred per household member, rather than to have households of very different sizes receiving the same transfer amount. To accomplish this, we scaled the transfer amounts within a village by household size, such that larger households received larger transfers, but leaving the mean transfer amount at the village level unaffected. This formula first calculated the per-capita transfer for a village using household sizes and the desired average household transfer value. Second, it scaled household-level transfer amounts with household size, applying a minimum of 3 members and a maximum of 8 members, so as to achieve the intended mean transfer amount per household per village.

emerge from this table are that Gikuriro was successfully targeted at the poorer households even within the (relatively poor) eligible group, and that they met more success with households headed by younger individuals. Conditional on this, however, they were not differentially successful at reaching households with children, or female-headed households. In the case of GiveDirectly's cash transfers, we see evidence only that GD's strict enforcement of government eligiblity criteria makes membership in official poverty groups a strong correlate of transfer receipt.

2.5 Cost Equivalence, Before and After the Fact

The design-based approach to cost-equivalent comparisons across programs is hampered by the fact that we can only measure costs well when the programs have been fully implemented, but the head-to-head comparison would need to know these costs at the design phase. Anticipating this issue, we costed both programs in detail prior to and after the intervention period following Levin and McEwan (2001), and also randomized cash transfer treatment amounts. The ex-ante costing exercise arrived at an ex-ante cost of $119 per beneficiary household. We then randomized transfer amounts at the village level to bracket this anticipated cost. Three smaller cash transfer costing $77, $119, and $152 (with beneficiaries actually receiving $41, $84, and $117, respectively) are used to form the cost-adjusted comparison. The larger cash transfer cost $567 and transferred $517.

The ex-post costing was conducted using the 'ingredients method', valuing inputs at their opportunity costs (see Dhaliwal and Tulloch, 2012; Levin and McEwan, 2001, for more discussion). The costing question is asked from the perspective of the donor (in this case, USAID), meaning that we consider the total money spent per beneficiary to achieve the benefits observed. Overhead expenditures in the implementation chain are an inherent part of these costs, and so the lower transactions costs in getting mobile money to the beneficiary play an important role in their potential attractiveness.[22]

Since the Gikuriro program covers eight districts (e.g. much larger than the study population only) and many of the startup and administrative costs must be amortized over this whole beneficiary

[22] Costs are inclusive of all direct costs, all indirect in-country management costs including transport, real estate, utilities, and the staffing required to manage the program, and all international overhead costs entailed in managing the Gikuriro program. All administrative costs, including the appropriate share of the costs of maintaining international headquarters infrastructure, were included in the costing. Beneficiary identification costs, incurred by the survey firm and identical across all arms of the study, are excluded from the cost-benefit calculation. Monitoring and Evaluation costs, similarly, were excluded so as to be costing only the implementation component.

pool, we calculate cost per beneficiary in the full national program. Because we did not want differences in scale to drive differential costs per beneficiary, we asked GiveDirectly to artificially scale up their operations and provide us with numbers reflecting the costs per beneficiary if they were running a national-scale program across eight districts, including 56,127 beneficiary households like Gikuriro. We costed each GD arm separately, asking what the overhead rate would have been if GD had run a national program at the scale of Gikuriro giving only transfers of that amount.[23] This allows us to conduct the benefit/cost comparisons 'at scale', rather than having the artificial, multi-amount environment of the study contaminate the costing exercise across arms.

The ex-post costing exercise arrived at a cost per beneficiary household for Gikuriro of $141.84. Actual GD costs per beneficiary were $66, $111, and $145, and $567. The numbers actually used for our cost-equivalence exercise, however, need to reflect the fact that compliance with cash was higher than compliance with Gikuriro. Since the regressions estimate the Intention to Treat (ignoring non-compliance), the relevant cost number is not the amount spent per beneficiary, but rather the effective spend per eligible household included in the ITT. To calculate this number, we differentiate 'averted' costs which are not incurred by the implementer in the case of non-compliance (for example costs of running a SILC in which the beneficiary does not participate) versus 'non-averted' costs (such as trainings that must be paid for no matter how many people attend). We then calculate the cost per eligible household as the sum of the non-averted costs per beneficiary and the averted costs multiplied times the compliance rate. For GD all costs are averted in the case of non-compliance, and for Gikuriro roughly 60% of costs (excepting WASH and BCC) are averted. Using this approach we can recover a cost-equivalent comparison even when the compliance rates are different across arms.

Table 2 provides the final costing numbers arrived at by the ex-post exercise with the costs per beneficiary in the top row, the averted share in the second row, compliance rates across arms in the third row, and the final cost per eligible study household used in the costing exercise in the bottom row. Gikuriro compliance rates are 80 percent among eligibles. Given an actual cost to USAID of $141.84 per beneficiary and the role of averted costs, this gives a cost of $124.49 per eligible household. GD compliance rates range from 81%–91% across arms. Given this, GD costs per

[23]Overhead costs as a percentage of the amount transferred decline sharply with transfer amount for GD because fixed costs represent a large share of their total overhead.

eligible household are \$54, \$96, \$121, and \$517, so the Gikuriro costs used in the Cost Equivalence exercise wind up being slightly higher than the largest of the three small arms that were intended to bracket the Gikuriro cost. We use these values as described in Section 4.1 and our pre-analysis plan to regression-adjust outcomes for the differential cost from Gikuriro, providing a way of testing the differential impact of the programs at identical cost to the donor.

2.6 Attrition and Balance of the Experiment

Attrition from follow-up is low at the household level, at around 3.3 percent in the control arm. Differences between arms are slight, ranging from 0.89 percentage points lower in the GD Main arm, to 1.7 percentage points lower in the GD Large arm; only the latter is significantly different from Control, and then only at the 10 percent level. As we show in Appendix C, patterns of attrition are not meaningfully predicted by household covariates. However, for individual-level anthropometric outcomes among children under 6, we see significantly lower attrition rates among children in treated households. Following our pre-analysis plan, we therefore also estimate models that use inverse probability weights to balance characteristics among the sample at follow up.

Next we present the baseline comparison of primary and secondary outcomes as well as household-level control variables, using the non-attrited panel sample that will be the basis of the endline evaluation. The regressions used here mimic as closely as possible the impact regressions, using fixed effects for randomization blocks, weighting to make the sample representative of all eligibles, and clustering standard errors at the village level to account for the design effect. All tables also present p-values adjusted for False Discovery Rates within outcome families (primary or secondary) using the technique of Anderson (2008); stars in the tables are based on clustered standard errors. Looking first at balance on study outcomes, Tables C.1, C.2 show balance for primary and secondary outcomes, respectively. Overall balance on secondary outcomes is excellent, but we do find evidence that we have superior outcomes on child anthropometrics in the GD Large arm. HAZ and WAZ are about 0.2 standard deviations higher than the control in this arm, and the latter difference is significant at the 10% level even after adjusting for multiple inference.[24] Per our pre-analysis plan, we present all core results on anthropometrics using an ANCOVA structure (controlling for baseline

[24] Note that this imbalance does not arise from attrition; if we estimate these differences using the full baseline sample the number of observations rises by 47 but the coefficients on GD Large remain very similar.

values) which should correct for this baseline imbalance. Table C.3 examines balance over baseline covariates and finds all treatment arms to be comparable to the control across all covariates.

3 Intention-to-Treat impacts

Tables 3 and 4 present the core intention-to-treat results of the study on primary and secondary outcomes for the eligible population. The analysis is conducted as an ANCOVA, including the lagged dependent variable where available, fixed effects for the randomization blocks, and for each outcome including a set of baseline controls selected by post-double-LASSO (Belloni et al., 2014b, as described in Online Appendix F) for it by LASSO as described in the Online Appendix. For parsimony, our main tables include a dummy for the Gikuriro treatment, another for the Large GD treatment, and then a single dummy that indicates the three smaller GD transfers. Standard errors are clustered at the village level to reflect the design effect, and we also present "sharpened" q-values that adjust inference to control the False Discovery Rate within each table (Anderson, 2008). Observations are weighted to make the analysis representative of all eligible study households, incorporating both survey weights arising from the number of eligible households per village, and the intensive tracking weights. Panel A of each table analyzes household-level outcomes, and Panel B individual-level outcomes.

Taking the Gikuriro treatment first, we see no statistically significant impacts on primary outcomes at the household level. Estimates are sufficiently precise to allow us to rule out impacts on consumption that would be sufficient to justify the program. For example, the upper bound of a 95 percent confidence interval for Gikuriro's impact on household consumption is 0.086, ruling out consumption gains of more than 9 percent (equivalent to 0.064 standard deviations of consumption). Neither are there significant impacts on household wealth (point estimate 0.01; upper bound of 95 percent CI 0.36 equivalent to 0.086 standard deviations), or, in secondary outcomes, measures of physical asset ownership. Dietary diversity, anthropometrics, and maternal anemia all move in the right direction but none of these changes is significant. 95 percent confidence intervals for HAZ and WAZ rule out impacts outside of the ranges of (-0.03, 0.13) and (-0.04, 0.12), respectively.[25]

[25]For comparison, these confidence intervals are sufficiently precise to rule out the estimated one-year HAZ impacts of 0.15 sd from a nutritional intervention from recent work in Bangladesh, but are essentially centered on the 0.05 sd one-year HAZ impacts of a combined nutrition and WASH intervention in the same study (Luby et al., 2018). A parallel trial in Kenya found its nutrition intervention to have a one-year impact of 0.11 sd, and its nutrition and

Examining impacts on outcomes more directly on Gikuriro's causal path, household savings increase by a 109 percent, consistent with the creation of SILCs, while no impacts appear for health knowledge or sanitation practices. Hence the program has been successful in moving an indicator closely related to one of its sub-components, but at least within the time-frame of the study these changes in savings have not translated into significant improvements in the core welfare outcomes.

We turn next to to the impact of the three smaller ("Main") GiveDirectly arms whose average cost per eligible is $90.28, 72 percent of the cost of Gikuriro. Transfers of this magnitude do not alter the primary outcomes, though we do see effects on secondary outcomes, in a manner quite distinct to Gikuriro's in-kind programming. For instance, point estimates of household consumption impacts of 0.06, and corresponding 95 percent confidence interval upper bound of 0.24, rule out impacts equivalent to 0.18 standard deviations or greater; we similarly rule out HAZ and WAZ impacts in excess of 0.06 and 0.07 standard deviations of the reference population, respectively. Instead of increasing savings, small GD transfers lead to a 77 percent pay-down of debt, and an increase in the value of productive and consumption assets, by 26 percent and 35 percent respectively. Thus far, then, the comparison of Gikuriro to cash breaks down into two distinct dimensions of improvement, each of which has a different and entirely plausible pathway to long-term improvements: savings (Gikuriro), or debt reduction and asset investment (GiveDirectly).[26]

When we examine the third column, however, a more transformative impact arising from the Large cash transfer is clearly apparent. Not only do omnibus measures of consumption and wealth go up across the board, but metrics of consumption closely linked to child health improve. The dietary diversity score increases by 0.55 food groups, or by 12 percent off a base of 4.77. Figure 2 displays the fraction of each arm consuming each food group, and shows the treatment effects of the large arm to be most pronounced in fish, fruits, oils and fats, spices and condiments, and cereals. Productive assets increase by 80 percent, consumer assets almost double in value, and home value increases substantially. In the individual outcomes the benefits of this surge in consumption are evident as well; within the course of one year we see a 0.09 SD improvement in HAZ, a 0.07 SD improvement in WAZ, and a 0.13 SD improvement in MUAC, all significant at least at the 10 percent level, prior to

WASH intervention to have a one-year impact of 0.12 standard deviations (Null et al., 2018).

[26]Table A.4 shows the experimental results with each of the three treatment cells within the 'main' treatment broken out individually. Because these cells are both small and feature treatment amounts that are similar, this more granular analysis does not turn up evidence of meaningful heterogeneity across the three smaller GD transfer amounts.

adjustment for multiple inference.[27] The ANCOVA specification should be particularly important in the analysis of the anthropometric indicators that showed signs of imbalance at baseline; indeed if we examine these outcomes in post-treatment levels we see substantially stronger apparent treatment effects. Appendix Table A.5 analyzes the binary anthropometric outcomes of stunting and wasting, and finds impacts very consonant with the continuous impacts on HAZ and WAZ. The GD large transfer reduces stunting by 6 pp on a base of 50% and wasting by 5 pp on a base of 16%, both effects significant at the 10% level. Anemia falls slightly (not significant) and there is a substantial decrease in child mortality of almost 1 percentage point(or 70 percent off of the baseline value). To contextualize these effects using unweighted numbers, the control group eligibles saw 13 cases of child mortality out of 2,596 children (0.5 percent) while the GD Large arm saw 2 cases out of 1,200 children (0.16 percent). Hence the GD Large arm saw meaningful improvements in consumption and child health.

In the final columns of each table we provide p-statistics on F-tests that the ratio of the benefits across arms differs from the ratio of their costs. These statistics ask whether we can reject that the impacts scale in a manner similar to the costs. For the comparison across the two GD arms we find two significant differences: only in the case of debt reduction (where small transfers have a big effect and big transfers do not) and house quality (where small transfers have a negative and large transfers a positive effect) can we reject cost-proportional benefit scaling for cash transfers. In comparing Gikuriro to the GD large arm we see more differences, with Gikuriro being more cost-effective in generating savings and GD Large superior in consumption, as well as productive and consumption assets.

Our pre-analysis plan states that for any outcomes where we find differential attrition, we would estimate a propensity to remain in the sample incorporating covariates, dummies for treatments, and their interactions on the right-hand side, and then re-weight the analysis by the product of the inverse of this probability and the standard sampling weight. This procedure corrects the impacts for the observable determinants of attrition, and uses regression weighting to attempt to make the treatment and control samples comparable on important covariates even after attrition. Because we primarily found significantly differential attrition for the anthropometric outcomes, in Table A.6

[27] These improvements should be viewed against the backdrop of a sharp deterioration in anthropometrics subsequent to birth that typically occurs in LDCs, leaving rural African children often two full SDs below the international norm by age 3 (Shrimpton et al., 2001; Victora et al., 2010).

we present the results of this correction. We interact with each treatment dummy the same right-hand side covariates we use the same controls in the anthropometric regressions: gender, a linear, quadratic, and cubic for age in months, baseline household wealth, and a dummy for membership in Ubudehe poverty category 1. The results in this table are virtually identical to Table 3, indicating that the types of children who attrited from the study are similar across arms and hence differential attrition is unlikely to be driving our impacts.[28]

In many ways the core surprise of these results are not those of the cash arms (which conform quite closely to the broader literature in terms of how transfer amounts translate into impact magnitudes), but that Gikuriro proves so ineffective even on its core outcomes. Why might this be the case? Three possibilities suggest themselves. First is the question of study timing, given that our endline occurs only 8-9 months after the implementation of Gikuriro. The two main comparison papers for this intervention (Leroy et al., 2016; Null et al., 2018) both focus on anthropometric measurement two years subsequent to treatment. Hence it is possible that given more time the Gikuriro intervention would have generated stronger impacts, but Null et al. (2018) show that the bulk of anthropometric gains arise within the first year,[29] and that adherence to the treatment decays substantially in the second year. Therefore, it is likely the case that we are studying the most intensely treated period of time, and that interventions which were going to impact child anthropometrics by year 2 would have had visible effects in Year 1. A second issue is that of cost. Gikuriro, as implemented, cost only $142 per beneficiary, while the highly effective Tubarumure program in neighboring Burundi cost between $676 and $766 dollars per beneficiary (Heckert et al., 2020). A program that devoted more resources to direct feeding and nutrition, in concert with the WASH components of Gikuriro, may have achieved stronger results on child anthropometrics.[30] A final

[28]An alternate approach to this differential attrition is the use of Lee Bounds (Lee, 2009); Tables A.7 and A.8 conduct a bounding exercise for the anthropometric outcomes and find that the impact of the Large cash arm on Height-for-Age remains significant at the 90% level even in the lower bound. The other anthropometric effect of this arm are not significant in the lower bound but still represent improvements of between 0.06 and 0.09 standard deviations.

[29]After one year, length-for age z scores have improved versus Control by 0.11 sd in their Nutrition arm and by 0.12 sd in their Nutrition and WASH arm, of an eventual 0.13 and 0.16 sd impact in Year 2, respectively (Null et al., 2018, Supplementary Appendix, Figure S5).

[30]The fact that the smaller cash transfers also had no significant effect on child growth is consistent with Davis and Handa (2015), who suggest that transfer sizes of at least 20 percent of baseline household expenditure are required to have broad impacts on such outcomes. In a recent meta-analysis of the cash transfers literature (Manley et al., 2022), transfers of average value comparable to our study (mean value USD 90) have only modest effects on child growth (HAZ effect 0.024, WAZ effect 0.19, both insignificant), with each increase in transfer value by 10% of household income raising this impact by an additional one percentage point. When cash transfers are conditional, transfer sizes may have an alternative pathway to health outcomes through strengthening incentives for compliance (De Groot et

concern could be implementation fidelity. Given, however, that we have independent verification of program compliance from the survey firm that achieves almost as high a takeup rate as the cash arm, this does not appear to be a concern in this case.

Finally, we investigate whether ITT effects exhibit statistically and economically meaningful heterogeneity by transfer modality or recipient characteristics. As shown in Appendix Tables A.13 and A.14 for primary and secondary outcomes, receipt via lump-sum is estimated to have meaningfully different effects on only a few outcomes, and these do not survive multiple hypothesis corrections. On the other hand, heterogeneity by recipient characteristics might imply that there are winners and losers among subgroups of the targeted population that are traded off in cash-versus-kind decisions; it also creates scope for finer targeting that might improve the average effects of intervention in the population. In Appendix E, we assess the extent of such predictable heterogeneity, testing for heterogeneous impacts by a set of pre-specified attributes that either define populations of interest to policymakers, or that are anticipated to be important for the relative efficacy of the treatment regimes studied. We find no statistically significant evidence of such differential effects, suggesting that fine targeting offers limited scope for improvements in impacts in this context.

4 Comparative cost effectiveness

4.1 Conceptual Framework

Having presented the ITT results from the study, we now introduce a conceptual framework to help structure an analysis of the tradeoffs across outcomes, across beneficiaries, and across expenditure scales that are inherent in comparisons of multiple programs. Consider a policymaker seeking to maximize, by choice of policy P, a utilitarian social welfare function of the form

$$W(P) = \sum_{i=1}^{N} \sum_{t=1}^{T} \delta^t \omega_i U(Y_{i,t}(P)), \tag{1}$$

al., 2017). For older children – aged above the 1,000 day window studied here – alternative modalities such as school feeding programs may deliver child growth benefits at lower costs (Afridi, 2010). Consequently, interventions seeking to address child malnutrition at low cost may be better served by directly addressing the source of malnutrition through widely tested interventions such as multiple micro-nutrient supplementation or complementary feeding programs for newborns (Keats et al., 2021)."

for utilities $U(\cdot)$ defined over a K-dimensional vector of outcomes for individual i at time t under policy P, $Y_{i,t}(P) = \{y_{1,t}(P), \ldots, y_{K,t}(P)\}$. The policymaker places values on individuals in population with weights ω and discounts the stream of utility this induces exponentially with rate δ. This representation of the policymaker's problem embodies additive separability of outcomes across individuals and time.

To operationalize this welfare function for the empirical evaluation of policies with potential effects across multiple outcomes, we further assume additive separability of individual utility in the outcome dimensions of $Y_{i,t}(P)$.[31] This allows us to write instantaneous utility as the weighted sum of component-wise utilities over outcome dimensions: $U(Y_{i,t}(P)) = \sum_{k=1}^{K} u_k(y_{k,t})$. Given additive separability, there exists a choice of units for each outcome y_k such that we can write $u_k(y_k) = \kappa_k y_k$, where $\kappa_k = \{\kappa_1, \ldots, \kappa_K\}$ are relative weights on outcome dimensions.

Using this formulation, we can compare a policymaker's preference over outcome distributions arising from policies A versus B. The welfare differential can then be written as:

$$W(A) - W(B) = \sum_{k=1}^{K} \kappa_k \sum_{i=1}^{N} \omega_i \left(y_{ik}(A) - y_{ik}(B) \right), \tag{2}$$

where $y_{ik}(P)$ is the potential outcome of individual i under policy regime $P \in \{A, B\}$. The outcome-specific components of this welfare differential can be estimated as the (ω-weighted) difference in average outcomes across policies.

Within this framework, we draw three observations.

First, policies will generally move the distribution of several outcomes, so that decisions between them require taking a stance on relative weights across outcomes, κ. Under additive separability and given a functional form choice for y such that $u_k(\cdot)$ is linear in y, an evaluation that estimates impacts across multiple dimensions can place bounds on the set of outcome weights, κ, that are required to rationalize a given policy choice.

Second, through general-equilibrium effects or other externalities (Egger et al., 2019; Haushofer et al., 2015), policy choices will typically affect both targeted beneficiaries and non-targeted individuals in a given population. The framework of equation (2) can accommodate anonymity only

[31] Since we will be interested in the changes in welfare arising from changes in policy P, this can be thought of as a first-order approximation.

within classes of targeted vs non-targeted individuals by allowing a relative weight of ω to be placed on all individuals in the non-targeted beneficiary group (with weights on targeted group serving as numeraire). A benchmarking approach that estimates average effects on these complementary targeted and non-targeted subpopulations will allow us to place bounds on the weights on non-targeted individuals required to support a preference for a given policy, P.

Third, when the costs of alternative policies are equivalent, then the welfare differential of equation (2) provides a sufficient statistic for the policymaker's decision criterion. But when the costs of policies A and B are different, it is no longer automatic that this represents a policy-relevant comparison. Policymakers with a fixed budget to allocate to these programs, faced with alternatives that differ in costs per beneficiary, can respond either on the extensive margin, by treating fewer individuals with the less expensive program, or on the intensive margin, by increasing some aspect of the 'dosage' of the less expensive program. Comparison in this case requires welfare impacts to scale linearly, either on the intensive or extensive margin, in program expenditure. Then, given welfare value of the control state, $W(0)$, and program costs $C(A)$, $C(B)$, expenditure of a given amount on policy A is preferred to the same expenditure on policy B if $\left(W(A) - W(0)\right)/C(A) - \left(W(B) - W(0)\right)/C(B) \geq 0$: that is, if it has a greater cost-effectiveness ratio.[32] Typical experimental comparisons of alternative programs will speak to these actually feasible at-scale outcome distributions only such additional, strong, and typically untested assumptions. Our cost-equivalence approach relaxes this linear-in-costs assumption and makes policy comparisons more transparent.

Given the randomization of cash transfer amounts in the neighborhood of the in-kind program, we can instead make a locally-linear assumption as to how the benefits of cash scale with costs. If the observed welfare under the in-kind program is $W(A)$, the cost of each cash arm j relative to the in-kind arm is $\tau_j \equiv C(B_j) - C(A)$, we can linearize the welfare under the cash arms around this cost as $W(B(Cost_j)) = \gamma W(B(\tau_j))$. By estimating the slope relationship between outcomes and expenditure in the cash arm, γ, in a neighborhood around the cost of the in-kind arm, and evaluating the cash-vs-kind difference at this point at which $\tau = 0$, we can then calculate the cost-equivalent welfare benefit of the in-kind program as $W(A) - \hat{\gamma} * W(B(\tau_j = 0))$. This is the

[32] If outcome weights κ are normalized to be equal to the marginal utility of an extra dollar in income, and if costs $C(P)$ are considered to be net of fiscal externalities, then this is equivalent to a comparison of the *marginal value of public funds* across programs, as in (Hendren and Sprung-Keyser, 2019).

difference between the observed welfare in the in-kind arm and the simulated welfare of cash had we observed the exact cost-equivalent transfer. If impacts are continuous in transfer values, this estimate of the cost-equivalent comparison can be made arbitrarily close to the truth by evaluating cash impacts sufficiently close to the in-kind costs. Moreover, comparisons based on alternative sets of cash values can be used to assess the robustness of any such estimate.

4.2 Cost-Equivalent Benchmarking

To operationalize our cost adjustment strategy, let the subscript i indicate the individual (or household, depending on unit of analysis), c the cluster (village), and b the randomization block. Outcomes are observed both at baseline (Y_{icb1}) and at endline (Y_{icb2}). First, begin with the total GD donor cost per eligible within each transfer amount arm, denoted by t_c. Subtract from this number the benchmarked Gikuriro cost per eligible household C described above, and denote the difference $t_c - C = \tau_c$; this is the deviation (positive or negative) of each GD arm from the benchmarked Gikuriro cost (all cost numbers are those given in the bottom row of Table 2. Set τ_c to zero in the control and Gikuriro arms. We can then run an ANCOVA impact regression as above, but controlling for a linear term in τ_c, an indicator T_c for receipt of either cash or in-kind treatment, and an indicator T_c^{GK} for receiving Gikuriro. We estimate this specification, as illustrated in equation (3) below, on eligible households in our sample:

$$Y_{icb2} = \alpha_b + \delta^{GK} T_c^{GK} + \delta^T T_c + \gamma_1 \tau_c + \beta X_{icb1} + \rho Y_{icb1} + \epsilon_{icb2}. \quad (3)$$

Because τ_c absorbs the deviation of the GD arm from the benchmarked Gikuriro cost, the coefficient on Gikuriro treatment, δ^{GK}, will provide a direct benchmarking test: this estimates the *differential* impact of Gikuriro benchmarked against an exactly donor-cost-equivalent cash transfer. In addition, subject to the assumption of linear transfer amount effects, the slope coefficient τ_c captures impacts arising from deviations in GD cost from Gikuriro cost, and the coefficient δ^T estimates the impact of GD at the cost of Gikuriro.

A graphical representation of our strategy is provided in the left-hand panel of Figure 1, which plots the average outcome on the y-axis for all four GD treatment amounts (colored circles), for GK (gray diamond), and the control (gray circle). The line represents the fitted average outcome by

GD transfer amount. By predicting the outcome on this line at the exact cost of Gikuriro (hollow circle), the benchmarked differential is then the vertical difference between the Gikuriro impact and the projected cost-equivalent GD impact.

The results of the cost equivalent analysis for primary and secondary outcomes are presented in Tables 5 and 6, respectively.

The first column of these tables contains the heart of the comparative benchmarking exercise (δ^{GK}), comparing the effects of Gikuriro's in-kind programming to cash at exact donor-cost-equivalent levels. There are well powered, economically important, and statistically significant differences between the interventions at cost-equivalent levels (even accounting for multiple inference) on one primary and six secondary outcomes. Gikuriro is significantly less effective than a cost-equivalent cash transfer at driving consumption levels, producing approximately 19 percent lower consumption. Standard errors of 0.08 on the Gikuriro differential in consumption imply power to detect impacts substantially smaller than the value of even the smaller cash grants. At cost-equivalent levels, we find no statistically significant difference on other primary outcomes. These null findings are powered to detect economically meaningful differences: for instance, the 95 percent confidence interval for impacts on the household dietary diversity score rules out differences of a magnitude greater than approximately 0.25 in either direction—an increase of one quarter of a food type on average, a seemingly reasonable aspiration for an intervention seeking to change nutritional practices. Likewise, we find no statistically significant difference in child-growth outcomes between the interventions (95 percent confidence interval: -0.058, 0.098); by contrast, the addition of a nutrition component to a water, sanitation, and hand-washing intervention in Kenya produced a difference of 0.17 in weight-for-age z-scores (Null et al., 2018).

We find statistically and economically significant differences in impacts at cost-equivalent levels across a range of secondary outcomes. Gikuriro is less effective at driving the paydown of debt and the accumulation of assets, while the in-kind program is significantly more effective than cash at creating savings. The differential effect of the programs on savings and borrowing is interesting, and suggests that while both interventions serve to improve the net stock of liquid wealth (savings net of borrowing), the focus on savings through the SILC groups in Gikuriro drives liquidity to be built through deposits, while households making their own choices are more strongly disposed to reduce debt instead. Which of these strategies makes more sense? A simple comparison of interest rates is revealing. Gikuriro SILCs were free to set their own interest rates, but typically paid about 5 percent per annum nominal. Credit interest rates, by comparison, vary from an average of 22 percent in the MFI sector to upwards of 60 percent in informal credit markets. Given that 32 percent of eligible households reported having both borrowing and savings at baseline (and 79 percent had either borrowing or saving) it seems that the desire to pay down debt might be warranted.

The second column of Tables 5 and 6 presents estimated coefficients on an indicator for assignment to any treatment (coefficient δ^T in equation (3)). This intercept term estimates the impact of cash transfers at a cost equivalent to Gikuriro, although this precise amount was not included in the experiment. Given that the average transfer amount across the three smaller cash arms is only slightly lower than the GK cost, this estimate looks generally similar to the second column of the ITT tables (the simple average experimental effect across those three transfer amounts). At the exact cost of Gikuriro, we estimate that cash transfers would have led to a significant 73 percent decrease in the stock of debt, and a 30 percent and 40 percent increase in productive and consumption assets, respectively.

The third column of Tables 5 and 6 provides a direct estimate of γ_1, the marginal effect of an additional 100 dollars in donor cost on primary and secondary outcomes, respectively. As could be inferred from ITT estimates, this coefficient is strongly significant across a wide range of outcomes, particularly those most related to household consumption. An extra 100 dollars leads to a 5 percent increase in consumption, a 9 percent increase in dietary diversity, a 19 percent increase in savings, an 13 percent and 14 percent increase in productive and consumption assets, respectively, and leads housing value to improve by 5 percent and the index of housing quality to increase by 0.1 SD. In terms of anthropometrics, the change in value of transfer is positive but small in absolute magnitude for HAZ, WAZ, and MUAC, and does not survive correction for multiple inference. An extra $100 per beneficiary household—with eligible households containing an average of 2.7 children under the age of six—increases HAZ by 0.022 standard deviations. Beyond this, none of the other individual outcomes respond to transfer amount in a manner that we can reject at 95 percent significance.[33]

While we pre-specified the simple linear functional form for interpolation of cash-transfer impacts

[33] Per the critique of Muralidharan et al. (2019) on the use of factorial designs, we may wish to compare Gikuriro to only the Lump Sum or Flow cash arms rather than bundling the two cash arms together. Tables A.9 and A.10 conduct the cost-equivalent analysis comparing to Lump Sum and Flow only, respectively, and find qualitatively similar results to Table 5.

to preserve statistical power, a natural question is the extent to which our benchmarking results are sensitive to the linear interpolation of cash-transfer impacts. To interrogate this, Tables A.11 and A.12 present seven different ways of forming the cost-equivalent comparison. Column 1 in these tables repeats the linear specification used elsewhere, column 2 uses a quadratic, and column 3 a cubic function in project cost, and the remaining columns use the linear specification but drop one of the cash transfer arms in each column. The table reports only the differential parameter of Gikuriro over cost-equivalent cash reported in Column 1 of Tables 5 and 6, and the associated standard error. Overall, the results prove highly robust to specification; all of the outcomes significant in the main specification are significant in at least five out of six of the remaining specifications, except for the differential on the sanitation practices index. These results also confirm the power gains arising from linear interpolation: standard errors for the differential effect of Gikuriro are generally substantially smaller in the linear specification than in quadratic or cubic specifications. To give a sense of magnitudes, to obtain a reduction in the variance of the cost-equivalent comparison for the consumption outcome equivalent to that arising from the linear specification, a researcher using a cubic specification would have to increase the sample size by 158 percent. Although our transfer amounts (with three closely bunched together and one much larger) do not provide a great deal of power to test for non-linearity, this exercise suggests that our core results are robust to alternate ways of forming the cost equivalence comparison.[34]

4.3 Cost Equivalence versus Cost Effectiveness

This study is designed to make a specific form of cost-equivalent comparison, namely the impact of a cash transfer intervention assessed at the exact cost of an in-kind intervention. This comparison fixes the amount to be spent per beneficiary and asks which intervention is more effective. A different but related question is that of cost effectiveness, where we compare programs that operate at different costs and ask which generates the greatest benefit per dollar spent—potentially making comparisons between programs that operate at radically different levels of resource intensity. Fixed costs and indivisibility in program design mean that cost-benefit ratios do not represent alternatives that can be delivered for a given budget to a given population, but instead represent possible gains

[34] One new result that emerges from these tables is that in some specifications Gikuriro is superior to cost-equivalent cash at improving HAZ, significant at the 10% level in three out of seven specifications.

to an additive welfare function for policymakers who are indifferent to the size of the beneficiary population (for a given budget) or are willing to adjust their budget (for a given population).[35]

The comparison between these two approaches is represented graphically in Figure 1 for the primary study outcomes. In the left-hand panel we plot comparative cost-equivalence; here the focus is on the value on the X-axis that is the ex-post cost of Gikuriro, and we are interested in seeing which is greater, the observed benefit of Gikuriro (represented by a black diamond) or the predicted benefit of cash at this cost (the hollow circle). In the right-hand panel we connect the shaded circle that represents the outcome in the control group (at zero cost) with the outcome in each arm; because the plot represents outcomes in benefit/cost space, the intervention that features the steepest slope in this graph has the highest cost effectiveness. Interestingly, this graphic illustrates that while in general there are not substantial differences between Gikuriro and cash at benchmarked cost, because the smallest cash arm is so inexpensive while producing outcomes that are generally as good as (or better than) more expensive treatments, for four out of the five outcomes represented the smallest cash transfer has the highest cost effectiveness. The difference in linearized cost effectiveness across arms is tested statistically for all primary and secondary outcomes in the F-statistics in the final columns of Tables 3 and 4.

Table 7 provides a statistical analysis of the benefit-cost slope terms represented in the right-hand panel of Figure 1, pooling the three smaller GD arms together as done in the rest of the paper. It presents the ITT coefficient from the prior tables divided by the cost of the arm, and so gives the household improvement generated per \$100 spent through that modality, as well as the corresponding standard error on this BCR.[36] The final three columns of this Table 7 provide statistical tests of the difference in cost-effectiveness slopes and show how difficult it is to power a study to reject these; despite the relatively large sample size of this study, it is only for the consumption outcome that we are able to reject equality of CBRs across interventions, with the cash interventions outperforming Gikuriro. Between the two cash arms we are unable to reject equal BCRs for any outcome, indicating benefits that scale linearly with cash transfer amounts.

It is also worth clarifying the relationship between design choices and the relative statistical

[35]Under the linear specification of equation (3), the question of whether the benefits of cash scale proportionally to expenditure amounts to the hypothesis that $\delta^T = \tau_c C$. We fail to reject this null across all primary outcomes.

[36]Given that individual-level outcomes in this table represent average benefits for amount spent on the household, if we want the BCRs at the individual level we need to scale up the coefficients in this table by the average number of individuals per household for each outcome.

power of the cost equivalence and cost effectiveness comparisons. Much of our ability to estimate a cost slope on cash comes from the (very expensive) large cash arm. Some of the comparisons highlighted here require a statistically well-estimated slope term on expenditure, and some do not. The cost equivalent comparisons are primarily powered by having sufficient observations close to the cost-equivalent level. If the two programs are similar in cost then studies can be well-powered without having a cleanly estimated slope term on the amount of expenditure, since this slope is not important if little cost-adjustment is required (this is the case in our study, as demonstrated by the robustness of the cost equivalence adjustment to dropping of cash arms shown in Table A.11). As the average cost of the two programs differs by a larger amount, then the cost adjustment becomes increasingly important and hence a precisely estimated slope term is key to the power of cost equivalence comparisons. Cost effectiveness comparisons inherently involve testing for differential significance of slopes (now simply cost-benefit ratios that go through the origin by definition), and hence rejection is more likely when we compare expensive programs that had strongly significant overall effects to begin with. For these reasons, caution should be exercised in attempting to perform this type of comparative evaluation when program budgets are limited and overall impacts are likely to be muted.

5 Conclusion

This study uses a large-scale randomized experiment to pose a number of questions in comparative cost-effectiveness. Most centrally, we establish an approach to ask whether it is better to run complex multi-dimensional programs or simply to provide cash grants of equal ex-post cost, when these costs are not known with perfect certainty before the trial. We combine randomization of cash-transfer amounts with linear interpolation of their effects to make exact ex-post cost-equivalent comparisons, and demonstrate that, while point estimates do not generally depend on linearity, this approach delivers power gains equivalent to more than a doubling of sample size relative to alternative methods of interpolation. A simple theoretical framework provides structure to the assumptions required to make welfare comparisons across outcomes, beneficiary groups, and program costs.

Rwanda may be a particularly interesting environment in which to pose the benchmarking question for several reasons. First, child malnutrition rates overall are high—the prevalence of stunting among children under age five in the 2014-15 Demographic and Health Survey was 37.9 percent, underweight 9.3 percent, and wasted 2.2 percent—though this represents an improvement in recent years (DHS, 2016) Second, Rwanda is a country notable in Africa for its bureaucratic capacity and the public health infrastructure has been successful in delivering substantial improvements in child and maternal health outcomes (NISR, 2015) through schemes such as Pay-for-Performance (Basinga et al., 2011). Hence, it may provide a relatively strong case in terms of interventions like Gikuriro that are led through the public health system and lean heavily on Community Health Workers (CHWs). Third, the Government of Rwanda has been experimenting extensively with cash transfer programs over the past few years, such as the inclusion of cash in the flagship *Umurenge* poverty reduction program (Gahamanyi and Kettlewell, 2015), the $50 million 'Cash-to-poor' program supported by the World Bank, as well as a number of efforts to transition the support systems for the country's large population of refugees to cash transfers (such as a World Food Programme (WFP) program that is now supporting 15,000 refugees in Gihembe Camp using cash rather than traditional in-kind aid mechanisms (Taylor et al., 2016)). Hence there should be the bureaucratic capacity to implement Gikuriro well, and there is both experience with and interest in cash transfers as a safety net modality in the country.[37]

Our application of this approach to the critical battle against child malnutrition highlights an even more basic lesson: evidently, the two intervention modalities studied here each require more than $140 per household to deliver clinically relevant impacts on child malnutrition outcomes within a year.[38] Interventions wishing to move child health at low cost may do better to focus on more direct approaches such as school feeding (Afridi, 2010) or micro-nutrient supplementation (Keats et al., 2021). Even at this relatively low cost, however, the programs do trigger meaningfully different responses in the household use of intertemporal assets, with cost-equivalent cash transfers leading to higher levels of consumption, the pay-down of debt, and growth in asset investment, while Gikuriro households save more than those receiving cost-equivalent cash transfers.

Our cost-equivalent estimates allow us to make the trade-off across outcomes very explicit. For example, a policy preference for Gikuriro is equivalent to asserting that the greater savings induced

[37]Given the framing provided by GiveDirectly and the unusually strong degree of social control exerted by local officials in the Rwandan context, it is certainly possible that our 'unconditional' transfers have been more forcibly devoted to child consumption than they would have been in a different context.

[38]Gikuriro's costs were substantially lower than a related program in neighboring Burundi.

by Gikuriro is worth foregoing the higher consumption flows and productive asset stocks obtained under a cost-equivalent cash transfer. In dollar terms, the estimates of Tables 5 and 6 mean that moving from the predicted mean consumption of a cost-equivalent cash transfer to Gikuriro implies giving up $28.32 in monthly consumption and adding $8.35 in debt stocks, in exchange for an increase of $18.70 in savings stocks. This kind of precision can help policymakers be much more exact as to the types of trade-offs required to justify one type of intervention over another.

A fundamental idea in development economics is that poor households should have a single "shadow value" of cash which pulls down investment in all capital-hungry endeavors in a symmetric way. Our ITT impacts are generally consistent with this view of the world, as an intervention that relaxes credit constraints leads to shifts in consumption patterns that are very broadly spread across domains. This property means that small cash transfers are hard to detect because they move too many outcomes by too small an amount to be significant, while large cash transfers result in a broad-based increase in consumption in many dimensions. It also means that benchmarking exercises that compare in-kind impacts against cash on a single-dimensional setting are likely to understate the welfare advantages of cash, even if stronger assumptions about preferences—as discussed in Section 4.1—are required to address this.

Rich survey sampling enables us to test how the program impacts differ according to the sub-populations targeted. Overall this analysis provides evidence of surprisingly homogeneous returns from these interventions, whether looking across pre-identified study subgroups, using machine learning to identify optimal targeting rules that span multiple outcomes, or by introducing self-targeting through beneficiary choice. The general lack of heterogeneity or scale effects, the lack of evidence that flow transfers are superior, absence of essential heterogeneity, and the uncorrelated individual-level benefits across different dimensions of impact all suggest that it is reasonable for policymakers to use simple lump-sum cash transfers and to target them using preference weights based on individual attributes or inequality.

Finally, comparisons across expenditure levels highlight a key policy tradeoff. When we compare the cost-equivalent programs to a cash transfer of almost five times the amount, we see that larger sums of money can not only powerfully improve overall consumption and dietary diversity, but also lead to modest improvements in child growth. Transfer amounts in this study are not well-powered to test the linearity assumption (three small transfers of similar size, one much larger), however we generally find outcomes scaling with transfer amount in a simple way. Randomized variation in transfer values lets us form a number of interesting counterfactuals; for example the smallest transfer at which the benefit of cash would exceed Gikuriro; for savings this number is $694, and for HAZ it is $277.[39] Policymakers seeking to move child growth outcomes face difficult tradeoffs between the depth and breadth of their interventions; the comparative ease with which the resource intensity of cash transfers can be adapted makes this counterfactual modality particularly capable of revealing these tradeoffs.

There are some important limitations to what can be learned in this application of cash benchmarking. First, the cross-village experiment lets us measure the impact of those parts of the Gikuriro intervention that are implemented at the village or household level only. Village-level assignment means that some of our tests involving transfer amounts are not highly powered statistically. Most importantly, the time frame of the study (13 months, but only 8–9 months after treatment with Gikuriro) means that we are measuring endline outcomes more quickly than would be ideal, particularly for anthropometrics that may respond slowly to improvements in nutrition. Gikuriro made substantial investments in local capacity around health and sanitation and this infrastructure may drive future benefits in a way not captured in our study. Both of the companion papers examining comparable programs have their core outcome measurement two years after implementation, a duration that we were not able to observe given the desire to implement Gikuriro universally (including to the control) as quickly as possible. Particularly given that we may expect cash effects to fade over time, understanding the relative time trajectory of treatment effects is an important agenda for future comparative evaluation work. However, given that the targeting criteria for these programs will cause children to age out of eligibility within a few years, it remains important to show that such programs can generate benefits quickly for a given set of vulnerable children, as do the larger cash transfers we study.

Given the nuance of our findings, it is hard to square them with any simple idea of cash transfers as a kind of uni-dimensional 'index fund'. While business investment may have a single, cardinal objective—financial profit—development policy is undertaken with many goals in mind, and a per-

[39] Evidence from the sister study in Rwanda (McIntosh and Zeitlin, 2022), which featured more intermediate treatment amounts, did in fact find evidence of a non-linear effect, with outcomes involving investment and time use peaking at transfers a little over $400. Nonetheless, in this study for no primary outcome are we able to reject a constant benefit/cost ratio across different cash amounts.

fect reconciling of these competing benefits would require a clear statement of trade-offs in this multi-dimensional space. Perhaps a clearer way of expressing the counterfactual provided by unconditional cash is that it gives us a statement of the priorities that the beneficiaries themselves hold when credit constraints are relaxed, and thereby motivates us to be clearer about the logic underlying paternalistic development programs. While beneficiary decisions may not be 'optimal' in terms of long-term social welfare (for example, due to high discount rates, to self- or other-control problem, or to resource and information constraints), the impact of unconditional cash is nonetheless a powerful statement of the outcomes that the beneficiaries themselves want changed. For us to argue that a program is justified in using resources to drive outcomes different from the ones the beneficiaries would choose, we should have a clear reason why they fail to arrive at the welfare-maximizing outcome themselves. This is a view of benchmarking that quantifies tradeoffs rather than picking a winner.

References

Afridi, Farzana, "Child welfare programs and child nutrition: Evidence from a mandated school meal program in India," *Journal of development Economics*, 2010, *92* (2), 152–165.

Aguero, Jorge, Michael Carter, and Ingrid Woolard, "The impact of unconditional cash transfers on nutrition: The South African Child Support Grant," 2006.

Ahmed, AU, JF Hoddinott, S Roy, E Sraboni, WR Quabili, and A Margolies, "Which Kinds of Social Safety Net Transfers Work Best for the Ultra Poor in Bangladesh," *Operation and Impacts of the Transfer Modality Research Initiative*, 2016.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney, "The long-run impact of cash transfers to poor families," *The American Economic Review*, 2016, *106* (4), 935–971.

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

Araujo, M Caridad, Mariano Bosch, and Norbert Schady, "Can Cash Transfers Help Households Escape an Inter-Generational Poverty Trap?," in "The Economics of Poverty Traps," University of Chicago Press, 2017.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or condition? Evidence from a cash transfer experiment," *The Quarterly Journal of Economics*, 2011, *126* (4), 1709–1753.

_ , _ , and _ , "When the Money Runs Out: Do Cash Transfers Have Sustained Effects on Human Capital Accumulation?," *World Bank Policy Research Working Paper*, 2016, *7901.*

_ , _ , and Berk Özler, "When the money runs out: Do cash transfers have sustained effects on human capital accumulation?," *Journal of Development Economics*, 2019, *140*, 169 – 185.

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

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry, "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, 2015, *348* (6236).

_ , Rema Hanna, Benjamin A Olken, Elan Satriawan, and Sudarno Sumarto, "Food vs. food stamps: Evidence from an at-scale experiment in Indonesia," NBER Working Paper no. 28641 April 2021.

Barham, Tania, Karen Macours, and John A Maluccio, "Schooling, Learning, and Earnings: Effects of a 3-Year Conditional Cash Transfer Program in Nicaragua After 10 Years," *La Plata, Argentina: Centro para los Estudios Distributivos, de Trabajo y Sociales. Disponible en: http://www. cedlas-er. org/sites/default/files/aux_files/barham-maluccio-macours_ . pdf*, 2014.

Basinga, Paulin, Paul J Gertler, Agnes Binagwaho, Agnes LB Soucat, Jennifer Sturdy, and Christel MJ Vermeersch, "Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation," *The Lancet*, 2011, *377* (9775), 1421–1428.

Belloni, A, D Chen, V Chernozhukov, and C Hansen, "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, November 2012, *80* (6), 2369–2429.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-dimensional methods and inference on structural and treatment effects," *The Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

_ , _ , and _ , "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 2014, *81*, 608–650.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen, "Turning a shove into a nudge? A labeled cash transfer for education," *American Economic Journal: Economic Policy*, 2015, *7* (3), 86–125.

Blattman, Christopher and Paul Niehaus, "Show them the money: why giving cash helps alleviate poverty," *Foreign Aff.*, 2014, *93*, 117.

_ , Nathan Fiala, and Sebastian Martinez, "The Long-Term Impacts of Grants on Poverty: Nine-Year Evidence from Uganda's Youth Opportunities Program," *American Economic Review: Insights*, 2020, *2* (3), 287–304.

Brudevold-Newman, Andrew Peter, Maddalena Honorati, Pamela Jakiela, and Owen W Ozier, "A firm of one's own: experimental evidence on credit constraints and occupational choice," 2017.

Card, David and Alan B Krueger, "Time-series minimum-wage studies: a meta-analysis," *American Economic Review*, May 1995, *85* (2), 238–243.

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

Chowdhury, Reajul, Elliott Collins, Ethan Ligon, and Munshi Sulaiman, "Valuing Assets Provided to Low-Income Households in South Sudan," 2017.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jaychandran, "The price effects of cash versus in-kind transfers," *Review of Economic Studies*, 2019, *86* (1), 240–281.

Currie, Janet and Douglas Almond, "Human capital development before age five," *Handbook of labor economics*, 2011, *4*, 1315–1486.

Davis, Benjamin and Sudhanshu Handa, "How Much Do Programmes Pay?," *children*, 2015, *1000*, 750.

Desai, Amy, Laura E Smith, Mduduzi NN Mbuya, Ancikaria Chigumira, Dadirai Fundira, Naume V Tavengwa, Thokozile R Malaba, Florence D Majo, Jean H Humphrey, and Rebecca J Stoltzfus, "The SHINE Trial infant feeding intervention: pilot study of effects on maternal learning and infant diet quality in rural Zimbabwe," *Clinical Infectious Diseases*, 2015, *61* (suppl_7), S710–S715.

Dhaliwal, Iqbal and Caitlin Tulloch, "From research to policy: using evidence from impact evaluations to inform development policy," *Journal of Development Effectiveness*, 2012, *4* (4), 515–536.

DHS, "Rwanda Demographic and Health Survey, 2014–15," Technical Report, National Institute of Statistics of Rwanda, Minisry of Finance and Economic Planning, Ministry of Health, and the DHS Program, ICF International March 2016.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker, "General equilibrium effects of cash transfers: experimental evidence from Kenya," Technical Report, National Bureau of Economic Research 2019.

Fernald, Lia CH, Paul J Gertler, and Lynnette M Neufeld, "10-year effect of Oportunidades, Mexico's conditional cash transfer programme, on child growth, cognition, language, and behaviour: a longitudinal follow-up study," *The Lancet*, 2009, *374* (9706), 1997–2005.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth, "The effects of poor neonatal health on children's cognitive development," *The American Economic Review*, 2014, *104* (12), 3921–3955.

Filmer, Deon, Jed Friedman, Eeshani Kandpal, and Junko Onishi, "Cash Transfers, Food Prices, and Nutrition Impacts on Ineligible Children," *The Review of Economics and Statistics*, 05 2021, pp. 1–45.

Fink, Günther, Rachel Levenson, Sarah Tembo, and Peter C Rockers, "Home-and community-based growth monitoring to reduce early life growth faltering: an open-label, cluster-randomized controlled trial," *The American journal of clinical nutrition*, 2017, *106* (4), 1070–1077.

Finkelstein, Amy and Nathaniel Hendren, "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, 2020, *34* (4), 146–167.

Gahamanyi, Vincent and Andrew Kettlewell, "Evaluating graduation: insights from the Vision 2020 Umurenge Programme in Rwanda," *IDS Bulletin*, 2015, *46* (2), 48–63.

Gertler, Paul J, Sebastian W Martinez, and Marta Rubio-Codina, "Investing cash transfers to raise long-term living standards," *American Economic Journal: Applied Economics*, 2012, *4* (1), 164–192.

Groot, Richard De, Tia Palermo, Sudhanshu Handa, Luigi Peter Ragno, and Amber Peterman, "Cash transfers and child nutrition: pathways and impacts," *Development policy review: the journal of the Overseas Development Institute*, 2017, *35* (5), 621–643.

Habicht, Jean-Pierre and David L Pelletier, "The importance of context in choosing nutritional indicators," *The Journal of nutrition*, 1990, *120* (11 Suppl), 1519–1524.

Haushofer, Johannes and Jeremy Shapiro, "The short-term impact of unconditional cash transfers to the poor: Experimental Evidence from Kenya," *The Quarterly Journal of Economics*, 2016, *131* (4), 1973–2042.

_ , James Reisinger, and Jeremy Shapiro, "Your gain is my pain: Negative psychological externalities of cash transfers," Technical Report, Working Paper, Princeton University 2015.

Heckert, Jessica, Jef L Leroy, Deanna K Olney, Susan Richter, Elyse Iruhiriye, and Marie T Ruel, "The cost of improving nutritional outcomes through food-assisted maternal and child health and nutrition programmes in Burundi and Guatemala," *Maternal & child nutrition*, 2020, *16* (1), e12863.

Hendren, Nathaniel and Benjamin D Sprung-Keyser, "A unified welfare analysis of government policies," Technical Report, National Bureau of Economic Research 2019.

Hidrobo, Melissa, John Hoddinott, Amber Peterman, Amy Margolies, and Vanessa Moreira, "Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador," *Journal of Development Economics*, 2014, *107*, 144–156.

Hirvonen, Kalle and John Hoddinott, "Beneficiary Views on Cash and In-Kind Payments: Evidence from Ethiopia's Productive Safety Net Programme," 2020.

Hoddinott, John, Harold Alderman, Jere R Behrman, Lawrence Haddad, and Susan Horton, "The economic rationale for investing in stunting reduction," *Maternal & Child Nutrition*, 2013, *9* (S2), 69–82.

_ , Susanna Sandström, and Joanna Upton, "The impact of cash and food transfers: Evidence from a randomized intervention in Niger," 2014.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond, "Long-run impacts of childhood access to the safety net," *The American Economic Review*, 2016, *106* (4), 903–934.

Keats, Emily C, Jai K Das, Rehana A Salam, Zohra S Lassi, Aamer Imdad, Robert E Black, and Zulfiqar A Bhutta, "Effective interventions to address maternal and child malnutrition: an update of the evidence," *The Lancet Child & Adolescent Health*, 2021, *5* (5), 367–384.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Leroy, Jef L, Deanna Olney, and Marie Ruel, "Tubaramure, a Food-Assisted Integrated Health and Nutrition Program in Burundi, Increases Maternal and Child Hemoglobin Concentrations and Reduces Anemia: A Theory-Based Cluster-Randomized Controlled Intervention Trial–3," *The Journal of nutrition*, 2016, *146* (8), 1601–1608.

_ , Paola Gadsden, Sonia Rodríguez-Ramírez, and Teresa González De Cossío, "Cash and in-kind transfers in poor rural communities in Mexico increase household fruit, vegetable, and micronutrient consumption but also lead to excess energy consumption," *The Journal of nutrition*, 2010, *140* (3), 612–617.

Levin, Henry M and Patrick J McEwan, *Cost-effectiveness analysis: Methods and applications*, Vol. 4, Sage, 2001.

Luby, Stephen P, Mahjbubur Rahman, Benjamin F Arnold, Leanne Unicomb, Sania Ashraf, Peter J Winch, Christine P Stewart, Farzana Begum ad Faruqe Hussain, Jade Benjamin-Chung, Elli Leontsini, Abu M Naser, Sarker M Parvez, Alan E Hubbard, Audrie Lin, Fosiul A Nizame, Kaniz Jannat, Ayse ERcumen, Pavani K Ram, Kishor K Das, Jaynal Abedin, Thomas F Clasen, Kathryn G Dewey, LIa C Fernald, Clair Null, Tahmeed Ahmex, and John M Colford, "Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised trial," *Lancet Global Health*, 2018, *6*, e302–315.

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein, "The impact of improving nutrition during early childhood on education among Guatemalan adults," *The Economic Journal*, 2009, *119* (537), 734–763.

Manley, James, Harold Alderman, and Ugo Gentilini, "More evidence on cash transfers and child nutritional outcomes: a systematic review and meta-analysis," *BMJ Global Health*, 2022, *7* (4), e008233.

_ **, Seth Gitter, and Vanya Slavchevska**, "How effective are cash transfers at improving nutritional status?," *World development*, 2013, *48*, 133–155.

McIntosh, Craig and Andrew Zeitlin, "Using household grants to benchmark the cost effectiveness of a USAID workforce readiness program," *Journal of Development Economics*, June 2022, *157*, 102875.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "One-time transfers of cash or capital have long-lasting effects on microenterprises in Sri Lanka," *Science*, 2012, *335* (6071), 962–966.

Molina-Millan, Teresa, Tania Barham, Karen Macours, John A Maluccio, and Marco Stampini, "Long-Term Impacts of Conditional Cash Transfers in Latin America: Review of the Evidence," Technical Report, Inter-American Development Bank 2016.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich, "Factorial designs, model selection, and (incorrect) inference in randomized experiments," Technical Report, National Bureau of Economic Research 2019.

NISR, "Rwanda Demographic and Health Survey 2014-15. Rockvill," Technical Report, National Institute of Statistics of Rwanda (NISR) [Rwanda], Ministry of Health (MOH) [Rwanda] Rockville, Maryland, USA. 2015.

Null, Clair, Christine P Stewart, Amy J Pickering, Holly N Dentz, Benjamin F Arnold, Charles D Arnold, Jade Benjamin-Chung, Thomas Clasen, Kathryn G Dewey, LIa C H Fernald, Alan E Hubbard, Patricia Kariger, Audrie Lin, Stephen P Luby, Andrew Mertens, Sammy M Njenga, Geoffrey Nyambane, Pavani K Ram, and John M Colford, "Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomized controlled trial," *Lancet Global Health*, 2018, *6*, e316–e329.

Pega, Frank, Stefan Walter, Sze Yan Liu, Roman Pabayo, Stefan K Lhachimi, and Ruhi Saith, "Unconditional cash transfers for reducing poverty and vulnerabilities: effect on use of health services and health outcomes in low-and middle-income countries," *The Cochrane Library*, 2014.

Ruel, Marie T, Harold Alderman, Maternal, Child Nutrition Study Group et al., "Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition?," *The Lancet*, 2013, *382* (9891), 536–551.

Samuels, Fiona and Maria Stavropoulou, "Being Able to Breathe Again: The Effects of Cash Transfer Programmes on Psychosocial Wellbeing," *The Journal of Development Studies*, 2016, *52* (8), 1099–1114.

Schwab, Benjamin et al., "In the form of bread? A randomized comparison of cash and food transfers in Yemen," in "Agricultural & Applied Economics Association 2013 AAEA & CAES Joint Annual Meeting" 2013, pp. 4–6.

Sedlmayr, Richard, Anuj Shah, Munshi Sulaiman et al., "Cash-Plus: Cash Transfer Extensions and Recipient Agency," Technical Report, Centre for the Study of African Economies, University of Oxford 2017.

Seidenfeld, David, Sudhanshu Handa, Gelson Tembo, Stanfield Michelo, Charlotte Harland Scott, and Leah Prencipe, "The impact of an unconditional cash transfer on food security and nutrition: the Zambia child grant programme," 2014.

Shrimpton, Roger, Cesar G Victora, Mercedes de Onis, Rosângela Costa Lima, Monika Blössner, and Graeme Clugston, "Worldwide timing of growth faltering: implications for nutritional interventions," *Pediatrics*, 2001, *107* (5), e75–e75.

Skoufias, Emmanuel, Susan W Parker, Jere R Behrman, and Carola Pessino, "Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in mexico [with comments]," *Economia*, 2001, *2* (1), 45–96.

Suri, Tavneet, "Mobile money," *Annual Review of Economics*, 2017, *9*, 497–520.

Taylor, J Edward, Mateusz J Filipski, Mohamad Alloush, Anubhab Gupta, Ruben Irvin Rojas Valdes, and Ernesto Gonzalez-Estrada, "Economic impact of refugees," *Proceedings of the National Academy of Sciences*, 2016, *113* (27), 7449–7453.

Victora, Cesar Gomes, Mercedes de Onis, Pedro Curi Hallal, Monika Blössner, and Roger Shrimpton, "Worldwide timing of growth faltering: revisiting implications for interventions," *Pediatrics*, 2010, pp. peds–2009.

Vivalt, Eva, "How much can we generalize from impact evaluations?," *Journal of the European Economic Association*, 2015.

Tables and Figures

Table 1: Research design

			GiveDirectly				
	Control	Gikuriro	Lower transfer ($66)	Middle transfer ($111)	Upper transfer ($145)	Large transfer ($566)	Total
Panel A. Village-level randomization							
Villages	74	74	22	22	22	34	248
Ineligible households	298	297	88	87	88	137	995
Eligible households	521	541	165	154	167	246	1,794
Panel B. Household-level randomization of cash payment modality among eligibles							
Flow transfers			83	87	104	147	421
Lump-sum transfers			51	50	41	68	210
Choice			31	17	22	31	101

Notes: Table gives the number of observations across the arms of the study. The first row is the number of villages, the unit of assignment, and the remaining rows are numb er of households.

Table 2: Intervention costs and compliance rates

		GiveDirectly			
	Gikuriro	Lower	Middle	Upper	Large
Cost to USAID per beneficiary	$141.84	$66.02	$111.09	$145.43	$566.55
Share averted if untreated	0.60	1.00	1.00	1.00	1.00
Compliance rate among eligibles	0.80	0.81	0.86	0.83	0.91
Cost to USAID per eligible household	$124.49	$53.58	$95.86	$121.24	$517.44

Notes: Table gives costs for each of the study arms. The first row is the cost per beneficiary that emerged from the ex-post costing exercise. The second row gives the share of program costs averted among non-compliers, and the third row the compliance rate. Using these the fourth row provides the average amount spent per study eligible household. The fifth and sixth rows provide cost numbers for the average household in the village population.

Table 3: ITT estimates: Primary outcomes

	Gikuriro	GiveDirectly Main	GiveDirectly Large	Control Mean (SD)	Obs.	R^2	p-values: F-tests GK=GD	GD=GDL	GK=GDL
Panel A. Household outcomes									
Consumption†	−0.11	0.06	0.30***	10.69	1750	0.14	0.03	0.02	0.00
	(0.10)	(0.09)	(0.11)	(1.34)					
	[0.41]	[0.57]	[0.02]						
Household dietary diversity score	0.19	0.17	0.55***	4.77	1751	0.18	0.86	0.01	0.00
	(0.12)	(0.15)	(0.13)	(1.84)					
	[0.38]	[0.41]	[0.00]						
Household non-land wealth†	0.01	0.00	0.40	13.04	1751	0.22	0.98	0.16	0.15
	(0.18)	(0.21)	(0.28)	(4.24)					
	[0.60]	[0.60]	[0.38]						
Panel B. Individual outcomes									
Height-for-age	0.05	−0.02	0.09**	−1.97	2125	0.71	0.08	0.01	0.37
	(0.04)	(0.04)	(0.05)	(1.10)					
	[1.00]	[1.00]	[0.78]						
Weight-for-age	0.04	0.01	0.07*	−1.04	2104	0.68	0.49	0.14	0.46
	(0.04)	(0.03)	(0.04)	(0.98)					
	[1.00]	[1.00]	[0.78]						
Mid-upper arm circumference	0.02	−0.01	0.13*	−0.59	1629	0.51	0.66	0.10	0.17
	(0.06)	(0.07)	(0.08)	(0.95)					
	[1.00]	[1.00]	[0.78]						
Child anemia	0.00	0.02	−0.01	0.18	2372	0.07	0.45	0.43	0.74
	(0.02)	(0.02)	(0.04)	(0.39)					
	[1.00]	[1.00]	[1.00]						
Maternal anemia	−0.02	−0.00	−0.03	0.12	1581	0.11	0.33	0.34	0.90
	(0.03)	(0.03)	(0.03)	(0.33)					
	[1.00]	[1.00]	[1.00]						

Notes: Table presents Intention to Treat impacts on primary outcomes, with the three study arms presented in rows and the three smaller GiveDirectly transfers pooled into the 'Main' treatment. Regressions include but do not report the lagged dependent variable, fixed effects for randomization blocks, and a set of LASSO-selected baseline covariates, and are weighted be representative of the eligible population. Standard errors (in parentheses) are clustered at the household level to reflect the design effect. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines. Reported p-values in final three columns derived from F-tests of hypotheses that impacts are equal between Gikuriro and GD Main (GK=GD), between GD Main and Large transfer amounts (GD=GDL), and between Gikuriro and GD Large (GK=GDL).

Table 4: ITT estimates: Secondary outcomes

		GiveDirectly		Control			p-values: F-tests		
	Gikuriro	Main	Large	Mean (SD)	Obs.	R^2	GK=GD	GD=GDL	GK=GDL
Panel A. Household outcomes									
Stock of borrowing†	0.067	−0.765**	−0.341	7.39	1751	0.12	0.00	0.21	0.27
	(0.350)	(0.316)	(0.397)	(4.82)					
	[0.71]	[0.05]	[0.64]						
Stock of saving†	1.115***	−0.128	0.656**	5.88	1751	0.16	0.00	0.03	0.17
	(0.324)	(0.345)	(0.329)	(4.87)					
	[0.01]	[0.71]	[0.14]						
Health knowledge index	−0.076	0.153	0.075	2.89	1751	0.04	0.49	0.86	0.75
	(0.368)	(0.321)	(0.468)	(4.01)					
	[0.71]	[0.71]	[0.71]						
Sanitation practices index	−0.280	0.146	0.087	−0.68	1751	0.07	0.06	0.84	0.19
	(0.219)	(0.227)	(0.284)	(2.71)					
	[0.37]	[0.71]	[0.71]						
Productive assets†	0.020	0.257**	0.800***	11.22	1751	0.29	0.02	0.00	0.00
	(0.100)	(0.100)	(0.116)	(1.81)					
	[0.71]	[0.05]	[0.00]						
Consumption assets†	−0.367	0.354*	0.932***	8.70	1751	0.35	0.01	0.02	0.00
	(0.240)	(0.206)	(0.243)	(4.08)					
	[0.28]	[0.22]	[0.00]						
House value†	−0.023	−0.029	0.196***	13.81	1654	0.34	0.92	0.00	0.00
	(0.056)	(0.061)	(0.061)	(0.87)					
	[0.71]	[0.71]	[0.01]						
Housing quality index	−0.195	−0.217*	0.211	−0.17	1751	0.10	0.89	0.03	0.03
	(0.146)	(0.132)	(0.174)	(1.46)					
	[0.36]	[0.24]	[0.37]						
Panel B. Individual outcomes									
Child mortality	−0.006	−0.004	−0.009**	0.01	2687	0.01	0.63	0.31	0.48
	(0.005)	(0.006)	(0.004)	(0.11)					
	[1.00]	[1.00]	[1.00]						
Pregnancy	−0.031	−0.035	−0.007	0.20	2552	0.08	0.90	0.34	0.32
	(0.025)	(0.031)	(0.027)	(0.40)					
	[1.00]	[1.00]	[1.00]						
Live birth	0.103	0.091	−0.068	0.68	411	0.13	0.87	0.04	0.03
	(0.079)	(0.072)	(0.081)	(0.47)					
	[1.00]	[1.00]	[1.00]						
Birth in facility	−0.046	0.069	−0.062	0.84	293	0.16	0.03	0.18	0.88
	(0.059)	(0.052)	(0.099)	(0.37)					
	[1.00]	[1.00]	[1.00]						
Any vaccinations in past year	0.010	−0.010	−0.005	0.72	1291	0.26	0.54	0.90	0.71
	(0.034)	(0.032)	(0.039)	(0.45)					
	[1.00]	[1.00]	[1.00]						
Completed vaccination schedule	0.011	−0.013	0.006	0.58	1291	0.17	0.52	0.63	0.90
	(0.039)	(0.035)	(0.038)	(0.49)					
	[1.00]	[1.00]	[1.00]						
Disease burden	−0.020	−0.031	−0.018	0.54	2680	0.05	0.74	0.72	0.97
	(0.032)	(0.030)	(0.033)	(0.50)					
	[1.00]	[1.00]	[1.00]						
Diarrheal prevalence	−0.003	−0.000	−0.007	0.09	2680	0.04	0.83	0.69	0.84
	(0.015)	(0.016)	(0.015)	(0.29)					
	[1.00]	[1.00]	[1.00]						

Notes: Table presents Intention to Treat impacts on secondary outcomes, with the three study arms presented in rows and the three smaller GiveDirectly transfers pooled into the 'Main' treatment. Regressions include but do not report the lagged dependent variable, fixed effects for randomization blocks, and a set of LASSO-selected baseline covariates, and are weighted be representative of the eligible population. Standard errors (in parentheses) are clustered at the household level to reflect the design effect. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines. Reported p-values in final three columns derived from F-tests of hypotheses that impacts are equal between Gikuriro and GD Main (GK=GD), between GD Main and Large transfer amounts (GD=GDL), and between Gikuriro and GD Large (GK=GDL).

Table 5: Cost-equivalent comparisons: Primary outcomes

	Gikuriro: Differential	Cost-equivalent GD impact	Transfer Cost	Control Mean	Observations	R^2
A. Household outcomes						
Consumption†	−0.19**	0.08	0.05**	10.69	1750	0.14
	(0.08)	(0.09)	(0.02)			
	[0.06]	[0.49]	[0.06]			
Household dietary diversity score	−0.01	0.20	0.09**	4.77	1751	0.18
	(0.13)	(0.14)	(0.03)			
	[0.71]	[0.28]	[0.06]			
Household non-land wealth†	−0.03	0.04	0.09	13.04	1751	0.22
	(0.20)	(0.21)	(0.07)			
	[0.71]	[0.71]	[0.30]			
B. Individual outcomes						
Height-for-age	0.06	−0.01	0.02**	−1.97	2125	0.71
	(0.04)	(0.04)	(0.01)			
	[0.96]	[1.00]	[0.40]			
Weight-for-age	0.02	0.02	0.01	−1.04	2104	0.68
	(0.04)	(0.03)	(0.01)			
	[1.00]	[1.00]	[0.96]			
Mid-upper arm circumference	0.02	0.00	0.03	−0.59	1629	0.51
	(0.06)	(0.06)	(0.02)			
	[1.00]	[1.00]	[0.96]			
Child anemia	−0.02	0.02	−0.01	0.22	2372	0.07
	(0.03)	(0.02)	(0.01)			
	[1.00]	[1.00]	[1.00]			
Maternal anemia	−0.02	−0.00	−0.00	0.12	1581	0.10
	(0.02)	(0.03)	(0.01)			
	[1.00]	[1.00]	[1.00]			

Notes: First column is a dummy for Gikuriro treatment, giving the differential effect of Gikuriro over cash at equivalent cost. Second column is a dummy for either treatment, giving the impact of cash at the cost of Gikuriro. Third column is the cost slope, measured as the dollar-value deviation (in hundreds of dollars) of the treatment received from the cost of Gikuriro. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table 6: Cost-equivalent comparisons: Secondary outcomes

	Gikuriro: Differential	Cost-equivalent GD impact	Transfer Cost	Control Mean	Observations	R^2
A. Household outcomes						
Stock of borrowing†	0.793***	−0.726**	0.094	7.39	1751	0.12
	(0.277)	(0.309)	(0.077)			
	[0.01]	[0.04]	[0.19]			
Stock of saving†	1.184***	−0.070	0.188**	5.88	1751	0.16
	(0.333)	(0.327)	(0.080)			
	[0.01]	[0.54]	[0.04]			
Health knowledge index	−0.227	0.151	−0.023	2.89	1751	0.04
	(0.329)	(0.316)	(0.103)			
	[0.44]	[0.54]	[0.54]			
Sanitation practices index	−0.421**	0.141	−0.014	−0.68	1751	0.07
	(0.213)	(0.220)	(0.069)			
	[0.06]	[0.44]	[0.54]			
Productive assets†	−0.283***	0.302***	0.125***	11.22	1751	0.29
	(0.097)	(0.096)	(0.026)			
	[0.01]	[0.01]	[0.00]			
Consumption assets†	−0.764***	0.397**	0.139**	8.70	1751	0.35
	(0.256)	(0.199)	(0.058)			
	[0.01]	[0.06]	[0.04]			
House value†	−0.017	−0.007	0.049***	13.81	1654	0.34
	(0.051)	(0.059)	(0.014)			
	[0.54]	[0.55]	[0.01]			
Housing quality index	−0.008	−0.187	0.104**	−0.17	1751	0.10
	(0.151)	(0.126)	(0.048)			
	[0.56]	[0.12]	[0.05]			
B. Individual outcomes						
Child mortality	−0.002	−0.004	−0.001	0.01	2687	0.01
	(0.005)	(0.006)	(0.001)			
	[1.00]	[1.00]	[1.00]			
Pregnancy	0.002	−0.033	0.007	0.20	2552	0.08
	(0.027)	(0.029)	(0.007)			
	[1.00]	[1.00]	[1.00]			
Live birth	0.024	0.079	−0.037**	0.68	411	0.13
	(0.067)	(0.069)	(0.017)			
	[1.00]	[1.00]	[1.00]			
Birth in facility	−0.104**	0.058	−0.029	0.84	293	0.16
	(0.052)	(0.051)	(0.023)			
	[1.00]	[1.00]	[1.00]			
Any vaccinations in past year	0.018	−0.008	−0.000	0.72	1291	0.26
	(0.032)	(0.031)	(0.009)			
	[1.00]	[1.00]	[1.00]			
Completed vaccination schedule	0.021	−0.010	0.003	0.58	1291	0.17
	(0.036)	(0.033)	(0.009)			
	[1.00]	[1.00]	[1.00]			
Disease burden	0.012	−0.031	0.004	0.54	2680	0.05
	(0.033)	(0.029)	(0.008)			
	[1.00]	[1.00]	[1.00]			
Diarrheal prevalence	−0.003	−0.000	−0.002	0.09	2680	0.04
	(0.014)	(0.015)	(0.004)			
	[1.00]	[1.00]	[1.00]			

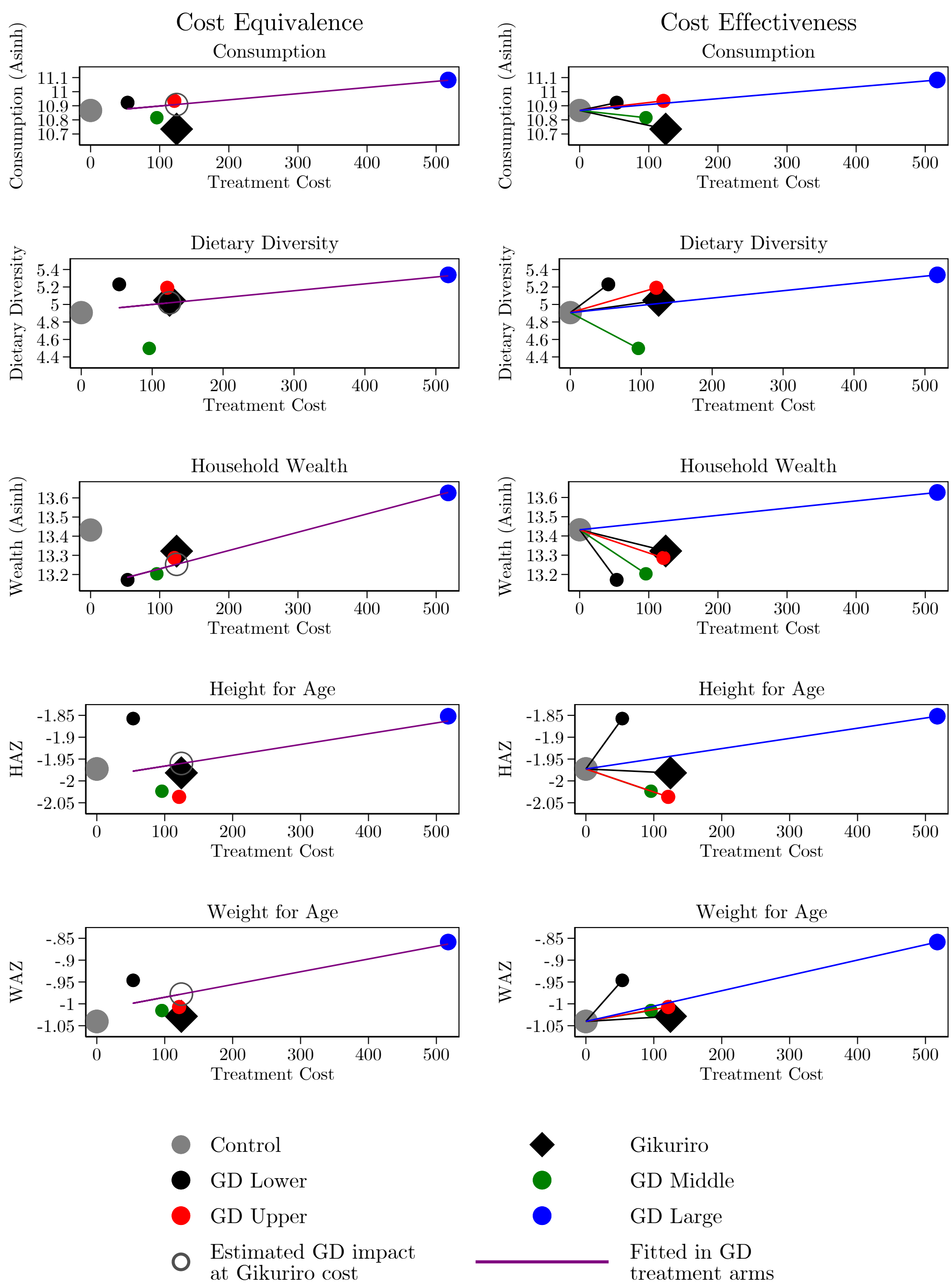
Notes: First column is a dummy for Gikuriro treatment, giving the differential effect of Gikuriro over cash at equivalent cost. Second column is a dummy for either treatment, giving the impact of cash at the cost of Gikuriro. Third column is the cost slope, measured as the dollar-value deviation (in hundreds of dollars) of the treatment received from the cost of Gikuriro. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table 7: Benefit-cost ratios for primary outcomes

		GiveDirectly		p-values		
	Gikuriro	Main	Large	(a)	(b)	(c)
Panel A. Household outcomes						
Consumption†	−0.090	0.066	0.058	0.05	0.93	0.04
	(0.077)	(0.100)	(0.021)			
Household dietary diversity score	0.155	0.186	0.106	0.83	0.60	0.59
	(0.098)	(0.162)	(0.024)			
Household non-land wealth†	0.007	0.005	0.077	0.99	0.74	0.60
	(0.148)	(0.235)	(0.054)			
Panel B. Individual outcomes						
Height-for-age	0.042	−0.021	0.018	0.11	0.33	0.47
	(0.036)	(0.044)	(0.009)			
Weight-for-age	0.031	0.012	0.013	0.61	0.97	0.54
	(0.032)	(0.037)	(0.007)			
Mid-upper arm circumference	0.018	−0.007	0.026	0.71	0.63	0.85
	(0.045)	(0.072)	(0.015)			
Child anemia	0.003	0.026	−0.002	0.38	0.30	0.77
	(0.018)	(0.028)	(0.007)			
Maternal anemia	−0.019	−0.002	−0.005	0.46	0.91	0.48
	(0.022)	(0.032)	(0.006)			

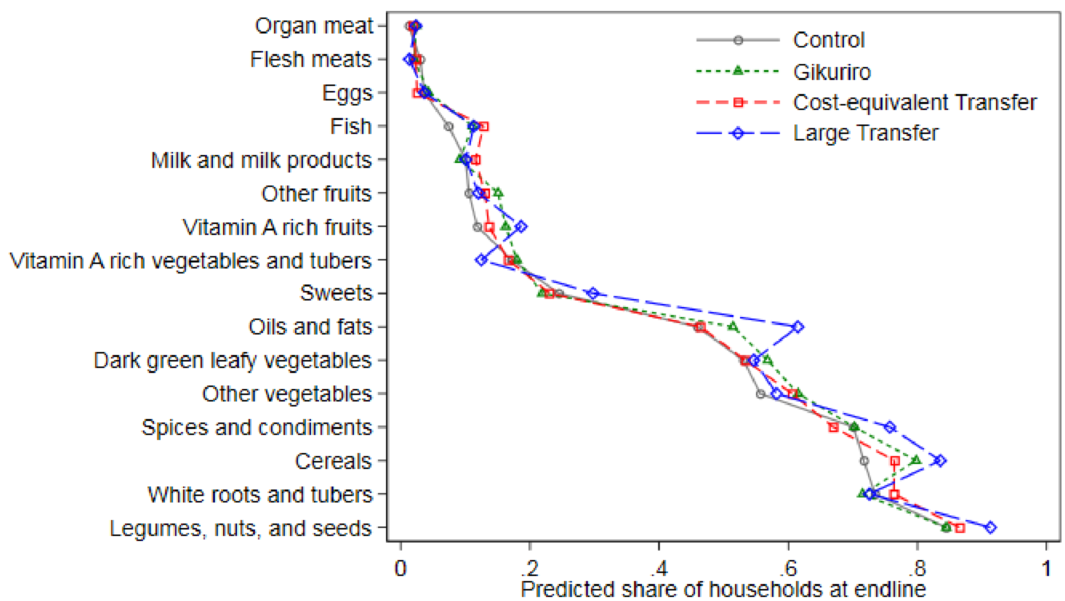
Notes: Benefit-cost ratios derived form ITT estimates and estimated costs per eligible household. *p*-values report tests of equal BCR between (a) Gikuriro and GD-Main; (b) GD-Main and GD-Large; and (c) Gikuriro and GD-large. Household-level BCRs; the average eligible household contains 5.2 members, 1.5 anthro-eligible children, 1.7 children eligible for anemia testing, and 1.2 adult women eligible for anemia testing. Per-person BCRs should be scaled up by these numbers.

Figure 1: Cost Equivalence versus Cost Effectiveness



Notes: Figures in left column visualize a cost-equivalent comparison (with no covariate adjustment). Dots represent mean outcomes in each treatment arm. Purple line represents a (population-weighted) regression of outcomes on treatment cost in the cash-transfer arms only. Hollow circle represents the point on that regression line for which expenditure per beneficiary is equivalent to the ex-post cost of Gikuriro; cost-equivalent comparison compares this to the diamond, which is the mean outcome in the Gikuriro arm. By contrast, figures in the right column illustrate a cost-benefit approach to comparing treatments (with no covariate adjustment). Here, the slope of the ray extending from the control group to the relevant treatment-specific mean represents the benefit-cost ratio; policymakers following this approach would favor the arm with the steepest slope.

Figure 2: Impacts on Dietary Diversity



Notes: Figure presents estimates of shares of households consuming goods that comprise the dietary diversity score in the past 7 days, by treatment arm. Control group represents unadjusted mean consumption rate of each food type. Other treatment arms represent the value in the control group, with the estimated impact of that treatment added, using the regression specification used to estimate intent-to-treat program impacts in Section **??**. Food types sorted by the share of control-group households who consume each item.

Online Appendices

Appendix A. Supplementary Tables and Figures

Table A.1: Correlates of compliance with intervention components

	Gikuriro					
	Nutrition Training	Cooking Training	Farmer Training	Farmer Harvest	Received Livestock	GD transfer
Number of Children	-0.0639	-0.0594	-0.0344	-0.0496	0.0351	-0.00711
	(0.0436)	(0.0390)	(0.0448)	(0.0551)	(0.0341)	(0.0341)
Number of Members	0.0815*	0.0894**	0.0679	0.0806	-0.0271	-0.00469
	(0.0423)	(0.0367)	(0.0409)	(0.0518)	(0.0301)	(0.0273)
Female headed HH	0.0190	-0.00581	-0.102	-0.0392	-0.0472	-0.0539
	(0.0691)	(0.0628)	(0.0762)	(0.0693)	(0.0650)	(0.0454)
Age of HH head	-0.00543***	-0.00436**	-0.00153	-0.00178	-0.00377**	-0.00109
	(0.00204)	(0.00190)	(0.00182)	(0.00182)	(0.00143)	(0.000969)
Poorest poverty group	0.150**	0.190**	0.222***	0.162**	0.238***	0.234**
	(0.0746)	(0.0766)	(0.0740)	(0.0756)	(0.0745)	(0.0973)
Next poorest group	0.165**	0.147**	0.151**	0.0869	0.211***	0.256***
	(0.0632)	(0.0589)	(0.0612)	(0.0687)	(0.0509)	(0.0839)
HH in Agriculture	0.0212	-0.0215	-0.0462	-0.0895	-0.0458	0.0285
	(0.0812)	(0.0762)	(0.0815)	(0.0808)	(0.0711)	(0.0448)
HH in Wage Work	0.0159	0.0539	0.0795	0.0994*	0.139***	0.0416
	(0.0584)	(0.0532)	(0.0483)	(0.0554)	(0.0487)	(0.0455)
HH in Microenterprise	-0.0460	0.0294	-0.0501	0.0194	-0.00755	0.0369
	(0.0688)	(0.0802)	(0.0604)	(0.0645)	(0.0662)	(0.0382)
Mean DV	0.628	0.511	0.568	0.480	0.335	0.863
N	524	524	524	529	529	718

Notes: Outcome is a dummy variable indicating that household received different components of Gikuriro (columns 1–5), or a cash transfer from GiveDirectly (column 6). First three columns are indicators for receiving training in proper nutrition (1), hygenic cooking habits (2), agricultural extension (3). Column 4 is an indicator that the household has successfully harvested a farm garden as instructed by FFLS, and Column 5 for the receipt of livestock from Gikuriro. Regressions are run among all households defined as eligible at baseline based on intended targeting criteria for the relevant treatment arm.

Table A.2: Attrition

	(1) Household	(2) Household	(3) Roster	(4) Roster	(5) Anthro	(6) Anthro	(7) Anemia	(8) Anemia	(9) New Member
Gikuriro	-0.013	-0.011	-0.010	-0.0086	-0.037***	-0.038***	-0.0044	-0.0046	-0.00050
	(0.0092)	(0.0085)	(0.0083)	(0.0079)	(0.013)	(0.012)	(0.021)	(0.016)	(0.0081)
GD Main	-0.0089	-0.0097	-0.0077	-0.0076	-0.024*	-0.025**	-0.0069	-0.0013	-0.0027
	(0.010)	(0.010)	(0.0093)	(0.0097)	(0.013)	(0.013)	(0.019)	(0.016)	(0.0088)
GD Large	-0.017*	-0.015	-0.015*	-0.013	-0.022	-0.020	0.027	0.0056	-0.0051
	(0.0097)	(0.0093)	(0.0086)	(0.0088)	(0.016)	(0.015)	(0.026)	(0.023)	(0.0096)
Agricultural		-0.015		-0.020		-0.037*		0.023	
		(0.012)		(0.013)		(0.020)		(0.023)	
Wage Worker		0.0076		0.016		0.019		0.0038	
		(0.0093)		(0.0099)		(0.015)		(0.016)	
Microenterprise		-0.0083		-0.0085		-0.00060		-0.039*	
		(0.0075)		(0.0063)		(0.017)		(0.022)	
Savings Group		0.00040		-0.0051		-0.015		-0.0020	
		(0.0085)		(0.0071)		(0.012)		(0.017)	
Village Eligibility Ratio		0.021		0.015		0.032		-0.0030	
		(0.053)		(0.048)		(0.050)		(0.060)	
Age of Head		-0.00034		-0.00025		0.00013		-0.0017***	
		(0.00022)		(0.00019)		(0.00034)		(0.00052)	
Schooling of Head		0.11		0.21		0.053		0.13***	
		(0.14)		(0.21)		(0.11)		(0.025)	
Dependency Ratio		0.035		0.027		0.058		0.029	
		(0.030)		(0.032)		(0.056)		(0.063)	
Household Size		-0.0036		-0.0031		-0.0060		-0.040***	
		(0.0023)		(0.0023)		(0.0040)		(0.0058)	
Poorest Category		0.0097		-0.0055		0.0065		-0.017	
		(0.011)		(0.0098)		(0.018)		(0.021)	
Next Poorest Category		0.0061		-0.0048		0.0059		0.0064	
		(0.0085)		(0.0094)		(0.014)		(0.017)	
female				0.0011		0.0012		-0.045***	
				(0.0026)		(0.0088)		(0.014)	
Control group mean	0.033	0.033	0.030	0.030	0.071	0.071	0.80	0.80	0.057
Observations	1793	1793	9189	9189	2265	2265	5192	5192	9509
R^2	0.0019	0.023	0.0014	0.028	0.0049	0.025	0.00068	0.068	0.000054

Notes: First two columns analyze household-level attrition from the baseline sample of households. (3) and (4) examine individual-level attrition from the roster of household members. (5) and (6) examine attrition from the baseline sample of child anthropometrics. (7) and (8) analyze attrition among individuals who should have received anemia testing, and (9) analyzes the probability of new household members appearing in rosters. Regressions are weighted to be representative of the eligible population.

Table A.3: Regressions with Granular GD Treatment Cells

		GiveDirectly				Control		
	Gikuriro	Small	Mid	Upper	Large	Mean	Obs.	R^2
Panel A. Household outcomes								
Consumption†	−0.11	0.10	−0.02	0.10	0.30***	10.69	1750	0.14
	(0.10)	(0.11)	(0.12)	(0.12)	(0.11)			
	[0.48]	[0.66]	[0.96]	[0.66]	[0.04]			
Household dietary diversity score	0.20	0.40**	−0.25	0.36**	0.56***	4.77	1751	0.19
	(0.12)	(0.19)	(0.22)	(0.15)	(0.13)			
	[0.29]	[0.12]	[0.48]	[0.10]	[0.00]			
Household non-land wealth†	0.01	0.06	0.04	−0.09	0.40	13.04	1751	0.22
	(0.19)	(0.40)	(0.26)	(0.26)	(0.28)			
	[0.96]	[0.96]	[0.96]	[0.96]	[0.41]			
Panel B. Individual outcomes								
Height-for-age	0.05	0.04	−0.04	−0.05	0.09**	−1.97	2125	0.71
	(0.04)	(0.05)	(0.06)	(0.05)	(0.05)			
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]			
Weight-for-age	0.04	0.01	−0.01	0.03	0.07*	−1.04	2104	0.68
	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)			
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]			
Mid-upper arm circumference	0.02	−0.04	0.07	−0.06	0.13*	−0.59	1629	0.51
	(0.06)	(0.08)	(0.12)	(0.07)	(0.08)			
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]			
Child anemia	0.00	0.05	0.01	0.02	-0.01	0.18	2372	0.07
	(0.02)	(0.06)	(0.03)	(0.02)	(0.04)			
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]			
Maternal anemia	−0.02	−0.03	−0.00	0.03	-0.03	0.12	1581	0.11
	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)			
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]			

Notes: Analysis includes dummies for each of the four GD transfer amounts separately. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines. Regressions are weighted to be representative of the eligible population.

Table A.4: Secondary Regressions with Granular GD Treatment Cells

		GiveDirectly				Control		
	Gikuriro	Small	Mid	Upper	Large	Mean	Obs.	R^2
Panel A. Household outcomes								
Stock of borrowing†	0.068	−0.670**	−0.780*	−0.846**	-0.339	7.39	1751	0.12
	(0.349)	(0.334)	(0.466)	(0.383)	(0.397)			
	[0.79]	[0.19]	[0.35]	[0.16]	[0.77]			
Stock of saving†	1.123***	−0.144	−0.584	0.353	0.661**	5.88	1751	0.16
	(0.326)	(0.422)	(0.487)	(0.507)	(0.329)			
	[0.01]	[0.79]	[0.49]	[0.77]	[0.19]			
Health knowledge index	−0.074	0.295	0.091	0.072	0.077	2.89	1751	0.04
	(0.369)	(0.480)	(0.410)	(0.378)	(0.468)			
	[0.79]	[0.79]	[0.79]	[0.79]	[0.79]			
Sanitation practices index	−0.273	0.299	−0.217	0.356	0.091	−0.68	1751	0.07
	(0.219)	(0.369)	(0.276)	(0.274)	(0.284)			
	[0.49]	[0.77]	[0.77]	[0.49]	[0.79]			
Productive assets†	0.021	0.270**	0.212	0.289**	0.801***	11.22	1751	0.29
	(0.100)	(0.122)	(0.151)	(0.138)	(0.116)			
	[0.79]	[0.16]	[0.49]	[0.19]	[0.00]			
Consumption assets†	−0.363	0.284	0.127	0.659**	0.933***	8.70	1751	0.35
	(0.241)	(0.337)	(0.267)	(0.266)	(0.244)			
	[0.45]	[0.77]	[0.79]	[0.11]	[0.00]			
House value†	−0.023	0.042	0.003	−0.130	0.196***	13.81	1654	0.34
	(0.056)	(0.079)	(0.087)	(0.086)	(0.061)			
	[0.79]	[0.79]	[0.87]	[0.45]	[0.01]			
Housing quality index	−0.195	−0.331	−0.283	−0.033	0.210	−0.17	1751	0.10
	(0.145)	(0.237)	(0.217)	(0.121)	(0.174)			
	[0.49]	[0.49]	[0.49]	[0.79]	[0.49]			

Notes: Analysis includes dummies for each of the four GD transfer amounts separately. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines. Regressions are weighted to be representative of the eligible population.

Table A.5: ITT Impacts on Stunting and Wasting

		GiveDirectly		Control			p-values: B/C ratios	
	Gikuriro	Main	Large	Mean	Obs.	R^2	GD=GDL	GK=GDL
Stunted	0.01	0.02	-0.06^{*}	0.50	2360	0.04	0.37	0.41
	(0.04)	(0.04)	(0.04)					
	[1.00]	[1.00]	[0.35]					
Wasted	0.00	-0.01	-0.05^{*}	0.16	2347	0.03	0.96	0.50
	(0.02)	(0.02)	(0.03)					
	[1.00]	[1.00]	[0.35]					

Notes: Table reports the Intention to Treat Impacts of the study arms on the binary outcomes of stunting and wasting (HAZ and WAZ respectively <-2). Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Variables marked with a † are in inverse hyperbolic sines. Regressions are weighted to be representative of the eligible population.

Table A.6: Anthropometric Impacts using Attrition IPW

	Gikuriro	Main	Large	Mean	Obs.	R^2
Height-for-age	0.051	−0.021	0.091**	−2.06	2125	0.71
	(0.045)	(0.039)	(0.046)			
	[0.62]	[1.00]	[0.35]			
Weight-for-age	0.038	0.010	0.067*	−1.06	2104	0.68
	(0.040)	(0.034)	(0.036)			
	[0.69]	[1.00]	[0.35]			
Mid-upper arm circumference	0.022	−0.007	0.135*	−0.58	1629	0.50
	(0.056)	(0.065)	(0.078)			
	[1.00]	[1.00]	[0.35]			

Notes: Regressions weighted using the product of standard survey weights and inverse propensity weights calculated from the probability that a child with baseline anthropometrics attrites from the endline.

Table A.7: Lower Lee Bounds for Anthropometric Effects.

	Gikuriro	Main	Large	Mean	Obs.	R^2
Height-for-age	0.010	−0.036	0.076*	−1.97	2080	0.70
	(0.042)	(0.039)	(0.045)			
	[1.00]	[1.00]	[0.92]			
Weight-for-age	0.004	−0.006	0.058	−1.04	2059	0.67
	(0.037)	(0.034)	(0.036)			
	[1.00]	[1.00]	[0.92]			
Mid-upper arm circumference	−0.037	−0.045	0.090	−0.59	1593	0.48
	(0.055)	(0.059)	(0.075)			
	[1.00]	[1.00]	[1.00]			

Notes: Table provides the Lower Lee Bounds for the anthro effects with the three treatment groups trimmed on the right tail of the outcome to match the (higher) degree of attrition observed in the control group. Regressions are weighted to be representative of the eligible population.

Table A.8: Upper Lee Bounds for Anthropomentric Effects.

	Gikuriro	Main	Large	Mean	Obs.	R^2
Height-for-age	0.071	0.001	0.107**	−1.97	2083	0.69
	(0.045)	(0.039)	(0.044)			
	[0.19]	[0.54]	[0.05]			
Weight-for-age	0.061	0.033	0.092**	−1.04	2062	0.67
	(0.041)	(0.034)	(0.036)			
	[0.19]	[0.32]	[0.05]			
Mid-upper arm circumference	0.065	0.029	0.181**	−0.59	1596	0.49
	(0.054)	(0.062)	(0.073)			
	[0.30]	[0.47]	[0.05]			

Notes: Table provides the Upper Lee Bounds for the anthro effects with the three treatment groups trimmed on the left tail of the outcome variable to match the (higher) degree of attrition observed in the control group. Regressions are weighted to be representative of the eligible population.

Table A.9: Cost-equivalent impacts of Gikuriro versus cash using only lump-sum cash transfers: Primary Outcomes.

	Gikuriro: Differential	Cost-equivalent GD impact	Transfer Cost	Control Mean	Observations	R^2
A. Household outcomes						
Consumption†	−0.22**	0.12	0.05	10.69	1379	0.15
	(0.09)	(0.10)	(0.03)			
	[0.21]	[0.60]	[0.50]			
Household dietary diversity score	0.05	0.16	0.08	4.77	1380	0.19
	(0.16)	(0.17)	(0.05)			
	[0.80]	[0.60]	[0.50]			
Household non-land wealth†	−0.21	0.22	0.07	13.04	1380	0.23
	(0.25)	(0.24)	(0.07)			
	[0.60]	[0.60]	[0.60]			
B. Individual outcomes						
Height-for-age	0.09**	−0.04	0.04***	−1.97	1663	0.72
	(0.04)	(0.04)	(0.01)			
	[0.33]	[1.00]	[0.04]			
Weight-for-age	0.03	0.01	0.02	−1.04	1641	0.70
	(0.05)	(0.04)	(0.01)			
	[1.00]	[1.00]	[1.00]			
Mid-upper arm circumference	−0.02	0.05	0.01	−0.59	1273	0.51
	(0.08)	(0.08)	(0.03)			
	[1.00]	[1.00]	[1.00]			
Child anemia	−0.02	0.03	−0.01	0.22	1852	0.08
	(0.03)	(0.03)	(0.01)			
	[1.00]	[1.00]	[1.00]			
Maternal anemia	−0.00	−0.02	−0.00	0.12	1238	0.10
	(0.03)	(0.03)	(0.01)			
	[1.00]	[1.00]	[1.00]			

Notes: Table estimated as in the main cost-effectiveness comparisons but uses only the Lump Sum cash treatments rather than pooling the Lump Sum and Flow cash arms together. Regressions are weighted to be representative of the eligible population.

Table A.10: Cost-equivalent impacts of Gikuriro versus cash using only "flow" cash transfers: Primary Outcomes.

	Gikuriro: Differential	Cost-equivalent GD impact	Transfer Cost	Control Mean	Observations	R^2
A. Household outcomes						
Consumption[†]	-0.19^{**}	0.07	0.05^{**}	10.69	1559	0.13
	(0.09)	(0.10)	(0.03)			
	[0.16]	[0.79]	[0.16]			
Household dietary diversity score	−0.03	0.22	0.09^{**}	4.77	1560	0.18
	(0.14)	(0.15)	(0.04)			
	[0.86]	[0.29]	[0.16]			
Household non-land wealth[†]	−0.12	0.11	0.03	13.04	1560	0.20
	(0.21)	(0.22)	(0.09)			
	[0.79]	[0.79]	[0.79]			
B. Individual outcomes						
Height-for-age	0.03	0.02	0.01	−1.97	1891	0.70
	(0.04)	(0.04)	(0.01)			
	[0.85]	[0.85]	[0.58]			
Weight-for-age	0.05	−0.02	0.01	−1.04	1874	0.68
	(0.04)	(0.04)	(0.01)			
	[0.58]	[0.85]	[0.58]			
Mid-upper arm circumference	0.11^{*}	-0.11^{*}	0.05^{**}	−0.59	1448	0.50
	(0.06)	(0.06)	(0.02)			
	[0.50]	[0.50]	[0.50]			
Child anemia	−0.03	0.04	−0.01	0.22	2114	0.08
	(0.03)	(0.03)	(0.01)			
	[0.58]	[0.58]	[0.58]			
Maternal anemia	−0.02	0.00	−0.01	0.12	1408	0.11
	(0.02)	(0.03)	(0.01)			
	[0.58]	[0.85]	[0.74]			

Notes: Table estimated as in the main cost-effectiveness comparisons but uses only the Flow cash treatments rather than pooling the Lump Sum and Flow cash arms together. Regressions are weighted to be representative of the eligible population.

Table A.11: Robustness of Linearity in Primary Cost Equivalence Adjustment

	Base Linear	Quad-ratic	Cubic	Drop lower	Drop mid	Drop upper	Drop huge
A. Household outcomes							
Consumption†	−0.193**	−0.164	−0.240*	−0.158*	−0.226***	−0.175**	−0.185*
	(0.079)	(0.102)	(0.127)	(0.086)	(0.086)	(0.084)	(0.103)
Household dietary diversity score	−0.010	0.107	−0.278	0.139	−0.208*	0.047	0.109
	(0.132)	(0.179)	(0.173)	(0.155)	(0.125)	(0.155)	(0.183)
Household non-land wealth†	−0.032	0.066	0.117	0.010	−0.028	−0.095	0.104
	(0.199)	(0.246)	(0.294)	(0.192)	(0.245)	(0.228)	(0.255)
B. Individual outcomes							
Height-for-age	0.060	0.098*	0.062	0.081*	0.030	0.047	0.095*
	(0.040)	(0.053)	(0.062)	(0.047)	(0.043)	(0.046)	(0.053)
Weight-for-age	0.024	0.009	−0.032	0.023	0.005	0.038	0.002
	(0.039)	(0.048)	(0.055)	(0.043)	(0.043)	(0.043)	(0.049)
Mid-upper arm circumference	0.017	0.037	0.108	0.018	0.055	−0.004	0.035
	(0.062)	(0.084)	(0.087)	(0.079)	(0.061)	(0.076)	(0.079)
Anemia	−0.018	−0.003	−0.016	−0.005	−0.027	−0.019	−0.002
	(0.026)	(0.025)	(0.029)	(0.021)	(0.032)	(0.031)	(0.026)
Anemia	−0.018	−0.052	−0.059	−0.035	−0.017	−0.003	−0.052
	(0.021)	(0.036)	(0.049)	(0.026)	(0.024)	(0.017)	(0.036)

Notes: Table reports the coefficient on the differential effect of Gikuriro over cost-equivalent cash using seven different specifications. Column 1 is the linear adjustment reported elsewhere. Column 2 includes a quadratic, and column 3 a quadratic and cubic term in the cost deviations from Gikuriro. Columns 4-7 leave out one of the cash treatment arms and repeat the linear cost adjustment. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Variables marked with a † are in inverse hyperbolic sines.

Table A.12: Robustness of Linearity in Secondary Cost Equivalence Adjustment

	Base Linear	Quad- ratic	Cubic	Drop lower	Drop mid	Drop upper	Drop huge
A. Household outcomes							
Stock of borrowing†	0.793***	0.917**	0.923**	0.847***	0.758***	0.767***	0.964***
	(0.277)	(0.360)	(0.411)	(0.322)	(0.278)	(0.293)	(0.371)
Stock of saving†	1.184***	1.071**	0.597	1.253***	0.958***	1.326***	1.037**
	(0.333)	(0.488)	(0.595)	(0.392)	(0.350)	(0.337)	(0.500)
Health knowledge index	−0.227	−0.112	−0.148	−0.140	−0.243	−0.285	−0.162
	(0.329)	(0.380)	(0.442)	(0.343)	(0.357)	(0.366)	(0.379)
Sanitation practices index	−0.421**	−0.397	−0.742**	−0.365	−0.569**	−0.352	−0.397
	(0.213)	(0.264)	(0.304)	(0.224)	(0.247)	(0.240)	(0.261)
Productive assets†	−0.283***	−0.241*	−0.283*	−0.246**	−0.309***	−0.287***	−0.218*
	(0.097)	(0.129)	(0.157)	(0.112)	(0.103)	(0.104)	(0.132)
Consumption assets†	−0.764***	−0.868***	−1.117***	−0.776***	−0.872***	−0.669**	−0.854***
	(0.256)	(0.304)	(0.337)	(0.275)	(0.276)	(0.279)	(0.304)
House value†	−0.017	0.082	0.128	0.029	−0.008	−0.069	0.086
	(0.051)	(0.073)	(0.092)	(0.059)	(0.057)	(0.054)	(0.074)
Housing quality index	−0.008	−0.112	−0.204	−0.043	−0.028	0.056	−0.189
	(0.151)	(0.168)	(0.162)	(0.160)	(0.161)	(0.175)	(0.163)
B. Individual outcomes							
Child mortality	−0.002	0.002	0.010**	−0.001	0.001	−0.006	0.003
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)	(0.006)
Pregnancy	0.002	−0.017	−0.052	0.000	−0.013	0.017	−0.023
	(0.027)	(0.041)	(0.047)	(0.035)	(0.027)	(0.029)	(0.041)
Live birth	0.024	0.011	−0.098	0.048	−0.021	0.080	−0.005
	(0.067)	(0.089)	(0.110)	(0.078)	(0.069)	(0.075)	(0.088)
Birth in facility	−0.104**	−0.141**	−0.177***	−0.122**	−0.134**	−0.060	−0.156***
	(0.052)	(0.061)	(0.068)	(0.058)	(0.055)	(0.063)	(0.060)
Any vaccinations in past year	0.018	0.060	0.089	0.039	0.025	−0.003	0.060
	(0.032)	(0.042)	(0.054)	(0.034)	(0.037)	(0.035)	(0.041)
Completed vaccination schedule	0.021	0.055	0.081	0.037	0.027	0.003	0.056
	(0.036)	(0.045)	(0.054)	(0.038)	(0.039)	(0.040)	(0.045)
Disease burden	0.012	−0.024	−0.034	−0.006	0.010	0.027	0.120***
	(0.033)	(0.040)	(0.045)	(0.035)	(0.036)	(0.036)	(0.039)
Diarrheal prevalence	−0.003	0.010	0.001	0.007	−0.010	−0.006	0.122***
	(0.014)	(0.023)	(0.032)	(0.017)	(0.017)	(0.014)	(0.038)

Notes: Table reports the coefficient on the differential effect of Gikuriro over cost-equivalent cash using seven different specifications. Column 1 is the linear adjustment reported elsewhere. Column 2 includes a quadratic, and column 3 a quadratic and cubic term in the cost deviations from Gikuriro. Columns 4-7 leave out one of the cash treatment arms and repeat the linear cost adjustment. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Variables marked with a † are in inverse hyperbolic sines.

Table A.13: Comparison of Lump Sum and Flow Transfers: Primary

	Main GD Treatment	Main GD Lump Sum	Large GD Treatment	Large GD Lump Sum	Control Mean	Observations	R^2
A. Household outcomes							
Consumption†	0.044	0.028	0.286***	0.027	10.69	1131	0.15
	(0.097)	(0.127)	(0.104)	(0.118)			
	[1.00]	[1.00]	[0.04]	[1.00]			
Household dietary diversity score	0.068	0.014	0.614***	−0.268	4.77	1131	0.19
	(0.159)	(0.176)	(0.137)	(0.237)			
	[1.00]	[1.00]	[0.00]	[1.00]			
Household non-land wealth†	0.012	−0.273	0.209	0.631	13.04	1131	0.25
	(0.262)	(0.502)	(0.388)	(0.441)			
	[1.00]	[1.00]	[1.00]	[1.00]			
B. Individual outcomes							
Height-for-age	−0.002	−0.095*	0.074	0.047	−1.97	1380	0.74
	(0.045)	(0.057)	(0.050)	(0.062)			
	[0.93]	[0.52]	[0.52]	[0.56]			
Weight-for-age	−0.036	0.125*	0.028	0.148**	−1.04	1369	0.70
	(0.040)	(0.071)	(0.041)	(0.073)			
	[0.56]	[0.52]	[0.56]	[0.52]			
Mid-upper arm circumference	−0.097	0.320***	0.121	0.150	−0.59	1057	0.53
	(0.063)	(0.112)	(0.082)	(0.107)			
	[0.52]	[0.11]	[0.52]	[0.52]			
Child anemia	0.044	−0.054	−0.007	−0.005	0.22	1544	0.08
	(0.032)	(0.038)	(0.039)	(0.052)			
	[0.52]	[0.52]	[0.92]	[0.93]			
Maternal anemia	0.013	−0.033	−0.024	0.009	0.12	1025	0.12
	(0.031)	(0.035)	(0.037)	(0.053)			
	[0.72]	[0.56]	[0.56]	[0.92]			

Notes: Analysis uses only the Control and GD arms; the first and third columns are dummies for GD transfer amount cells and the second and fourth are dummies measuring the additional effect of lump sum transfers within each cell. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table A.14: Comparison of Lump Sum and Flow Transfers: Secondary

	Main GD Treatment	Main GD Lump Sum	Large GD Treatment	Large GD Lump Sum	Control Mean	Observations	R^2
A. Household outcomes							
Stock of borrowing†	−1.080*** (0.323) [0.02]	1.071** (0.469) [0.16]	0.024 (0.411) [1.00]	−0.356 (0.892) [1.00]	7.39	1131	0.16
Stock of saving†	−0.584 (0.375) [0.51]	0.888** (0.429) [0.22]	0.573 (0.410) [0.59]	0.009 (0.723) [1.00]	5.88	1131	0.17
Health knowledge index	0.229 (0.360) [1.00]	0.115 (0.499) [1.00]	0.186 (0.554) [1.00]	0.176 (0.591) [1.00]	2.89	1131	0.05
Sanitation practices index	0.080 (0.302) [1.00]	0.406 (0.511) [1.00]	0.252 (0.290) [0.91]	0.040 (0.544) [1.00]	−0.68	1131	0.08
Productive assets†	0.199* (0.121) [0.47]	0.069 (0.295) [1.00]	0.792*** (0.135) [0.00]	0.148 (0.225) [1.00]	11.22	1131	0.29
Consumption assets†	0.336 (0.239) [0.59]	0.442 (0.352) [0.72]	0.831*** (0.300) [0.07]	0.465 (0.475) [0.90]	8.70	1131	0.38
House value†	0.023 (0.065) [1.00]	−0.162** (0.068) [0.16]	0.169* (0.088) [0.27]	0.012 (0.111) [1.00]	13.81	1071	0.35
Housing quality index	−0.260 (0.214) [0.72]	0.177 (0.433) [1.00]	0.206 (0.206) [0.90]	0.253 (0.235) [0.88]	−0.17	1131	0.11
B. Individual outcomes							
Child mortality	−0.001 (0.010) [1.00]	−0.002 (0.014) [1.00]	−0.010** (0.004) [0.60]	0.010 (0.009) [1.00]	0.01	1751	0.02
Pregnancy	−0.028 (0.032) [1.00]	−0.031 (0.034) [1.00]	−0.033 (0.029) [1.00]	0.045 (0.052) [1.00]	0.20	1646	0.09
Live birth	0.057 (0.088) [1.00]	0.099 (0.112) [1.00]	−0.087 (0.093) [1.00]	−0.123 (0.148) [1.00]	0.68	273	0.22
Birth in facility	0.077 (0.058) [1.00]	−0.013 (0.073) [1.00]	−0.095 (0.100) [1.00]	−0.033 (0.117) [1.00]	0.84	188	0.25
Any vaccinations in past year	−0.013 (0.039) [1.00]	0.038 (0.059) [1.00]	−0.022 (0.051) [1.00]	0.048 (0.101) [1.00]	0.72	838	0.27
Completed vaccination schedule	−0.006 (0.039) [1.00]	0.013 (0.070) [1.00]	−0.008 (0.050) [1.00]	0.063 (0.103) [1.00]	0.58	838	0.21
Disease burden	−0.016 (0.032) [1.00]	−0.055 (0.058) [1.00]	−0.004 (0.044) [1.00]	0.014 (0.084) [1.00]	0.54	1746	0.06
Diarrheal prevalence	−0.005 (0.018) [1.00]	0.021 (0.026) [1.00]	0.002 (0.019) [1.00]	−0.021 (0.024) [1.00]	0.09	1746	0.05

Notes: Analysis uses only the Control and GD arms; the first and third columns are dummies for GD transfer amount cells and the second and fourth are dummies measuring the additional effect of lump sum transfers within each cell. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table A.15: Effect of Transfer Modality Choice: Primary

	Got Ones Choice	Chose Lump Sum	Treated Lump Sum	Control Mean	Observations	R^2
A. Household outcomes						
Consumption†	0.089	−0.018	−0.032	10.69	534	0.18
	(0.115)	(0.126)	(0.117)			
	[1.00]	[1.00]	[1.00]			
Household dietary diversity score	0.067	0.062	−0.204	4.77	534	0.23
	(0.210)	(0.197)	(0.178)			
	[1.00]	[1.00]	[1.00]			
Household non-land wealth†	−0.650	0.033	0.172	13.04	534	0.29
	(0.468)	(0.471)	(0.460)			
	[1.00]	[1.00]	[1.00]			
B. Individual outcomes						
Height-for-age	−0.160**	0.034	−0.026	−1.97	671	0.75
	(0.062)	(0.058)	(0.063)			
	[0.21]	[1.00]	[1.00]			
Weight-for-age	−0.042	0.013	0.093*	−1.04	668	0.67
	(0.052)	(0.045)	(0.049)			
	[1.00]	[1.00]	[0.36]			
Mid-upper arm circumference	0.070	0.025	0.153**	−0.59	520	0.57
	(0.071)	(0.065)	(0.067)			
	[1.00]	[1.00]	[0.21]			
Child anemia	0.015	−0.017	−0.034	0.22	750	0.08
	(0.024)	(0.030)	(0.030)			
	[1.00]	[1.00]	[1.00]			
Maternal anemia	0.007	0.008	−0.004	0.12	496	0.14
	(0.034)	(0.043)	(0.037)			
	[1.00]	[1.00]	[1.00]			

Notes: Analysis uses only the GD arm. First column is a dummy for getting the chosen transfer modality, the second column is an (endogenous) indicator for choosing lump sum, and the third column is a dummy for actually receiving the lump sum treatment. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table A.16: Effect of Transfer Modality Choice: Secondary

	Got Ones Choice	Chose Lump Sum	Treated Lump Sum	Control Mean	Observations	R^2
A. Household outcomes						
Stock of borrowing†	−0.558	−0.197	0.956*	7.39	534	0.16
	(0.504)	(0.425)	(0.494)			
	[1.00]	[1.00]	[0.67]			
Stock of saving†	−0.222	0.074	0.132	5.88	534	0.21
	(0.463)	(0.491)	(0.450)			
	[1.00]	[1.00]	[1.00]			
Health knowledge index	0.056	0.559	0.524	2.89	534	0.09
	(0.534)	(0.479)	(0.454)			
	[1.00]	[1.00]	[1.00]			
Sanitation practices index	0.068	0.632*	0.221	−0.68	534	0.15
	(0.286)	(0.349)	(0.383)			
	[1.00]	[0.67]	[1.00]			
Productive assets†	−0.082	−0.166	0.001	11.22	534	0.29
	(0.212)	(0.207)	(0.203)			
	[1.00]	[1.00]	[1.00]			
Consumption assets†	−0.189	−0.917***	0.952***	8.70	534	0.37
	(0.362)	(0.319)	(0.301)			
	[1.00]	[0.06]	[0.05]			
House value†	−0.065	−0.029	−0.094	13.81	508	0.40
	(0.085)	(0.090)	(0.070)			
	[1.00]	[1.00]	[1.00]			
Housing quality index	−0.048	−0.017	−0.114	−0.17	534	0.21
	(0.175)	(0.119)	(0.123)			
	[1.00]	[1.00]	[1.00]			
B. Individual outcomes						
Child mortality	0.001	−0.011	0.000	0.01	838	0.03
	(0.010)	(0.014)	(0.015)			
	[1.00]	[1.00]	[1.00]			
Pregnancy	0.008	0.027	−0.025	0.20	757	0.11
	(0.035)	(0.032)	(0.030)			
	[1.00]	[1.00]	[1.00]			
Live birth	−0.064	0.286*	−0.007	0.68	129	0.32
	(0.151)	(0.162)	(0.120)			
	[1.00]	[1.00]	[1.00]			
Birth in facility	0.137	0.077	−0.153	0.84	83	0.57
	(0.127)	(0.121)	(0.115)			
	[1.00]	[1.00]	[1.00]			
Any vaccinations in past year	0.012	0.102*	0.016	0.72	434	0.25
	(0.050)	(0.062)	(0.058)			
	[1.00]	[1.00]	[1.00]			
Completed vaccination schedule	−0.004	0.060	0.002	0.58	434	0.19
	(0.054)	(0.065)	(0.064)			
	[1.00]	[1.00]	[1.00]			
Disease burden	0.017	0.050	−0.026	0.54	835	0.07
	(0.051)	(0.050)	(0.049)			
	[1.00]	[1.00]	[1.00]			
Diarrheal prevalence	−0.001	−0.030	0.017	0.09	835	0.05
	(0.024)	(0.026)	(0.025)			
	[1.00]	[1.00]	[1.00]			

Notes: Analysis uses only the GD arm. First column is a dummy for getting the chosen transfer modality, the second column is an (endogenous) indicator for choosing lump sum, and the third column is a dummy for actually receiving the lump sum treatment. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Regressions are weighted to be representative of the eligible population. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

Table A.17: Cash versus Kind Heterogeneity by Behavioral Attributes

	(1) Consumption†	(2) Diet Diversity	(3) Wealth†	(4) Borrowing†	(5) Saving†	(6) Health	(7) Sanitation	(8) Prod assets†	(9) Cons assets†
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Inconsistent x GK	0.22	0.37	-0.34	0.85	1.04	-0.34	0.54	0.56**	0.23
	(0.20)	(0.25)	(0.39)	(0.77)	(0.82)	(0.68)	(0.42)	(0.27)	(0.42)
Inconsistent x GD	0.42*	0.31	0.00021	0.25	0.69	0.40	0.68	0.56*	0.66
	(0.23)	(0.27)	(0.50)	(0.77)	(0.87)	(0.70)	(0.47)	(0.30)	(0.52)
Impatient x GK	-0.044	-0.49	0.025	-0.86	-0.57	0.78	-0.90*	-0.59**	-1.06*
	(0.20)	(0.32)	(0.52)	(0.85)	(0.87)	(0.78)	(0.53)	(0.26)	(0.57)
Impatient x GD	0.30	-0.037	-0.19	-0.29	-0.072	-0.020	0.21	0.11	-0.86*
	(0.22)	(0.32)	(0.52)	(1.02)	(0.83)	(0.90)	(0.47)	(0.33)	(0.48)
Lack Other Cont x GK	0.16	-0.33	-0.67*	0.89	-0.54	-1.51***	0.28	-0.32	-0.060
	(0.22)	(0.26)	(0.40)	(0.64)	(0.54)	(0.57)	(0.39)	(0.24)	(0.48)
Lack Other Cont x GD	0.16	-0.27	-0.63	-0.34	-0.44	-0.60	-0.11	-0.34	-0.23
	(0.22)	(0.28)	(0.45)	(0.68)	(0.65)	(0.73)	(0.43)	(0.29)	(0.52)
Time Inconsistent	-0.17	-0.22	-0.041	-0.41	-0.98	-0.19	-0.67**	-0.61***	-0.67**
	(0.17)	(0.20)	(0.25)	(0.61)	(0.70)	(0.42)	(0.27)	(0.17)	(0.29)
Impatient	-0.18	0.34	-0.10	0.088	0.62	-0.27	0.094	0.14	1.07***
	(0.16)	(0.27)	(0.33)	(0.72)	(0.62)	(0.60)	(0.38)	(0.20)	(0.36)
Lack Other Control	-0.16	0.050	0.25	0.20	1.01**	0.77*	-0.045	0.20	-0.16
	(0.16)	(0.19)	(0.21)	(0.47)	(0.41)	(0.39)	(0.27)	(0.19)	(0.27)
Gikuriro	-0.29*	0.16	0.43	-0.54	0.69	0.43	-0.51	-0.12	-0.31
	(0.16)	(0.21)	(0.33)	(0.63)	(0.60)	(0.68)	(0.32)	(0.19)	(0.32)
GiveDirectly	-0.33**	0.011	0.23	-0.81	-0.47	0.14	-0.39	-0.080	0.10
	(0.16)	(0.23)	(0.49)	(0.60)	(0.63)	(0.69)	(0.39)	(0.22)	(0.42)
Control Mean	10.4	4.16	12.9	5.96	5.18	0.19	-0.23	11.4	8.71
Observations	1508	1509	1509	1509	1509	1509	1509	1509	1509
R^2	0.15	0.18	0.22	0.12	0.17	0.049	0.079	0.32	0.38
Self_control_GD=GK	0.29	0.80	0.52	0.38	0.61	0.35	0.77	0.98	0.42
Other_control_GD=GK	0.97	0.85	0.95	0.066	0.88	0.22	0.38	0.93	0.77

Notes: Analysis excludes the GD Large arm, and uses interaction with three behavioral parameters to ask whether the effect of GD Main or Gikuriro are heterogeneous by Impatience, Time Inconsistency, or Other Control Problems. SEs clustered at the village level are in parentheses.

Table A.18: Lump Sum vs Flow Heterogeneity by Behavioral Attributes

	(1) Consumption†	(2) Diet Diversity	(3) Wealth†	(4) Borrowing†	(5) Saving†	(6) Health	(7) Sanitation	(8) Prod assets†	(9) Cons assets†
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Inconsistent x Lump Sum	0.37	0.34	0.74	0.097	0.48	0.061	1.10**	0.77**	0.96
	(0.25)	(0.30)	(0.70)	(1.05)	(0.91)	(0.97)	(0.48)	(0.33)	(0.69)
Inconsistent x Flow	0.35*	0.13	0.72	0.17	0.51	-0.48	0.58	1.13***	1.09*
	(0.19)	(0.30)	(0.61)	(0.89)	(0.80)	(0.86)	(0.38)	(0.27)	(0.61)
Impatient x Lump Sum	0.31	-0.27	0.27	-1.61	-0.58	0.047	0.47	0.49	-0.42
	(0.26)	(0.32)	(0.76)	(1.10)	(1.12)	(0.80)	(0.48)	(0.56)	(0.77)
Impatient x Flow	0.18	-0.40	-1.17*	1.23	-0.059	0.21	0.086	-0.41	-0.95
	(0.23)	(0.34)	(0.68)	(1.06)	(0.83)	(0.90)	(0.48)	(0.36)	(0.73)
Lack Other Cont x Lump Sum	0.28	0.11	-0.75	0.087	-0.27	-0.78	-0.68	0.093	0.41
	(0.21)	(0.33)	(0.88)	(1.08)	(0.71)	(0.75)	(0.71)	(0.34)	(0.80)
Lack Other Cont x Flow	0.19	-0.25	-0.028	0.076	0.15	0.33	-0.54	-0.21	-0.64
	(0.23)	(0.30)	(0.55)	(0.81)	(0.70)	(0.84)	(0.49)	(0.29)	(0.48)
Time Inconsistent	-0.13	-0.19	-0.25	-0.32	-0.72	0.22	-0.73***	-0.81***	-0.61*
	(0.15)	(0.18)	(0.23)	(0.55)	(0.64)	(0.42)	(0.23)	(0.18)	(0.36)
Impatient	-0.15	0.40*	-0.029	-0.064	0.53	-0.48	0.10	0.090	0.56
	(0.17)	(0.23)	(0.28)	(0.61)	(0.58)	(0.59)	(0.31)	(0.19)	(0.43)
Lack Other Control	-0.12	-0.067	0.082	0.054	0.86**	0.48	0.019	0.14	-0.13
	(0.16)	(0.18)	(0.21)	(0.43)	(0.43)	(0.36)	(0.27)	(0.18)	(0.26)
GD Lump Sum	-0.27	-0.0054	-0.35	0.35	0.54	0.34	-0.26	-0.20	0.15
	(0.20)	(0.29)	(0.69)	(0.82)	(0.72)	(0.76)	(0.51)	(0.29)	(0.61)
GD Flow	-0.21	0.47*	-0.28	-1.08*	0.050	-0.029	-0.14	-0.094	-0.0084
	(0.16)	(0.25)	(0.54)	(0.64)	(0.62)	(0.65)	(0.31)	(0.22)	(0.44)
Control Mean	10.4	4.16	12.9	5.96	5.18	0.19	-0.23	11.4	8.71
Observations	1131	1131	1131	1131	1131	1131	1131	1131	1131
R^2	0.15	0.19	0.25	0.16	0.18	0.060	0.091	0.31	0.38
Self_control_LS=Flow	0.93	0.60	0.97	0.95	0.97	0.61	0.31	0.27	0.85
Other_control_LS=Flow	0.65	0.31	0.49	0.99	0.58	0.26	0.85	0.40	0.26

Notes: Analysis uses only the control and the GD Main arm assigned to Lump Sum or Flow transfers to ask whether the impact of Lump Sum or FLow transfers is heterogeneous by Impatience, Time Inconsistency, or Other Control Problems. SEs clustered at the village level are in parentheses. .

Table A.19: Getting Flow when one Chose It, Heterogeneity by Behavioral Attributes

	(1) Consumption†	(2) Diet Diversity	(3) Wealth†	(4) Borrowing†	(5) Saving†	(6) Health	(7) Sanitation	(8) Prod assets†	(9) Cons assets†
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Inconsistent x Got It	0.61	0.075	1.93	-0.28	0.52	-2.14	-1.07	0.77	1.46
	(0.41)	(0.84)	(1.41)	(1.86)	(1.45)	(1.73)	(1.06)	(0.69)	(1.20)
Impatient x Got It	-0.32	-0.79	-3.60*	3.60	0.78	1.94	-0.056	-1.03	-2.85**
	(0.46)	(0.71)	(1.85)	(2.74)	(2.58)	(1.80)	(1.35)	(0.81)	(1.28)
Lack Other Cont x Got It	0.084	0.29	1.54	-0.86	0.53	-1.57	-1.46	0.12	-0.095
	(0.48)	(0.74)	(2.13)	(2.01)	(1.77)	(1.76)	(0.99)	(0.78)	(1.24)
Time Inconsistent	-0.38	0.18	-0.24	0.49	-0.039	1.76	0.80	-0.40	-0.43
	(0.36)	(0.79)	(0.71)	(1.33)	(1.30)	(1.30)	(0.91)	(0.61)	(0.91)
Impatient	0.42	0.30	2.78**	-2.92	-0.34	-2.75*	0.59	0.71	1.84*
	(0.41)	(0.62)	(1.37)	(2.21)	(2.40)	(1.50)	(1.14)	(0.71)	(1.02)
Lack Other Control	-0.22	-0.86	-2.81	-0.38	-0.95	2.40*	0.34	-0.53	-1.97**
	(0.37)	(0.59)	(2.03)	(1.66)	(1.63)	(1.42)	(0.87)	(0.72)	(0.97)
Got Choice in Choice Experiment	-0.41	-0.0070	-1.50	-1.37	-1.40	0.40	1.18	-0.75	-1.39
	(0.40)	(0.86)	(1.15)	(1.49)	(1.29)	(1.31)	(0.81)	(0.66)	(0.96)
Control Mean	10.4	4.16	12.9	5.96	5.18	0.19	-0.23	11.4	8.71
Observations	200	200	200	200	200	200	200	200	200
R^2	0.27	0.36	0.33	0.21	0.31	0.25	0.27	0.33	0.37

Notes: Analysis uses only the GD Main households that chose Flow transfers, and asks whether the impact of actually getting this choice (rather than being assigned to Lump Sum) is heterogeneous by Impatience, Time Inconsistency, or Other Control Problems. SEs clustered at the village level are in parentheses.

Figure A.1: Actual and Assigned Treatment Amounts

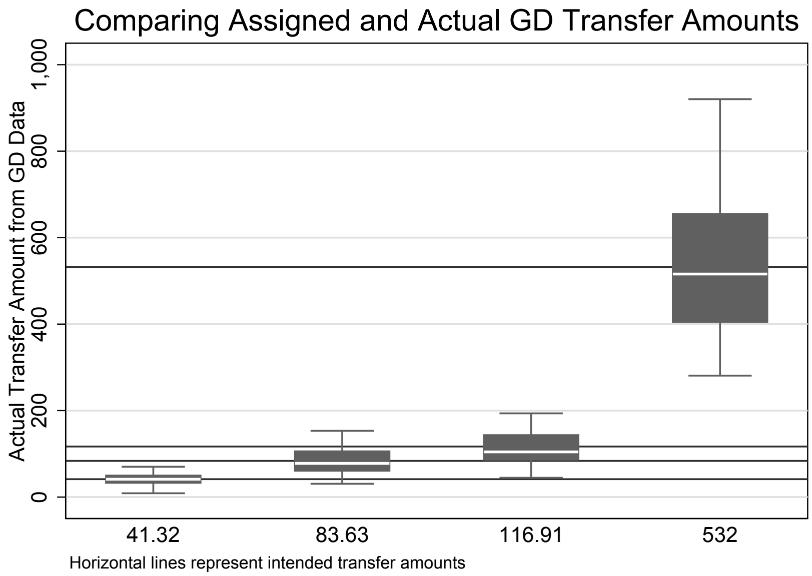


Figure A.2: GiveDirectly Nutritional Handout

USAID FROM THE AMERICAN PEOPLE Google.org GiveDirectly

GDID: ________________________ PAYMENT METHOD: ______________________

RECIPIENT INFORMATION REGARDING NUTRITION AND HYGIENE

GiveDirectly's program is supported by made possible by the generous support of the American people through USAID. The information below is approved by the Rwanda Ministry of Health.

- **Infant Nutrition**
  - Infants less than 6 months old should be fed by breast only. During this period an infant receives only breast milk and no other liquids or solids, not even water, unless medically indicated. A non-breastfed baby is 14 times more likely to die than an exclusively breastfed baby in the first 6 months.
  - Infants 6 to 24 months old should continue to be fed by breast, but should also receive complementary feeding that includes animal-source foods (meats, fish, milk products, eggs) and fruits and vegetables that are rich in vitamin A (such as mango, papaya, oranges, yellow sweet potato and carrots). Guidelines are for kids 6-24 months to eat at least 4 food groups: fruits, vegetables and legumes, grains, meats, dairy.
    - Infants 6 to 8 months old should be fed complementary foods 2-3 times daily;
    - Infants 9 to 24 months old should be fed complementary foods 3-4 times daily, plus 1-2 snacks.
- **Reducing Illness**
  - If you or your children get diarrhoea, use *Oral Rehydration Salts* (ORS) to replace the nutrients being lost. Typical symptoms of diarrhoea include frequent, loose, watery stools, abdominal cramps, and/or abdominal pain. If ORS is not available, a simple solution can be prepared for drinking by mixing one liter of clean drinking water and mix it with ½ teaspoon of salt and 6 teaspoons of sugar.
  - The government has a 6-monthly deworming program and Vitamin A supplementation program. Ask your Community Health Worker for more information.
- **Dietary Diversity**
  - Anemia
    - Anemia is a health condition, commonly caused by nutritional deficiency of iron and other nutrients (folate or vitamin B12). Around 72% of 6-8 months-olds in Rwanda have Anemia. Anemia can be an underlying cause for maternal death and prenatal and perinatal infant loss. Anemia among children is associated with low mental performance and physical development.
    - Examples of iron-rich food: fish, meat, milk products, oranges, lemons, grapefruits, guavas, papayas, and green leafy vegetables. Breast milk for your child is an important source of iron, too.
  - Here are some other examples of food you can produce/buy/eat to cheaply increase nutrition:

1

    - Breeding small, inexpensive animals such as hens, rabbits and guinea pigs can provide you and your children with important body building protein and other important nutrients.
    - Grow kitchen gardens if you have time. You can grow different vegetables for your family throughout the year, like amaranths, carrots, and dark-green leaves such as spinach and dodo, all of which are important sources of body protecting nutrients.
    - Consume soya beans, yogurt, avocados and dodo (which you could grow)
    - Eat orange-flesh rather than white-flesh sweet potatoes
- **Hygiene**
  - Handwashing with soap or wood ash can kill bacteria/viruses and prevents the spread of disease. Handwashing with soap at critical times is estimated to reduce diarrhoea by 47%. The most important times that hands should be washed with soap and water are:
    - After defecating
    - After cleaning a child who has defecated
    - Before eating or handling food
  - Recommended practices for personal hygiene further include:
    - Washing hair every week with shampoo
    - Washing the face every day after sleeping
    - Brushing teeth twice every day, in the morning and the night after eating
  - Safe disposal of waste means defecating into a latrine, disposing into a latrine, or burial. Inappropriate disposal of human feces, such as open defecation, facilitates the transmission of pathogens and disease.
- **Birth preparedness for delivery**
  - Early initiation of antenatal care (ANC) can reduce common maternal complications and maternal and perinatal mortality. Visit your nearest health facility early during pregnancy for medical tests and more information. The World Health Organization promotes four antenatal clinic visits, one in each trimester, during each pregnancy.

2

Appendix B. Eligibility for the Study

The study aims to compare nutrition and health gains among poor households with young children across the two programs and a control. We therefore used a definition of eligibility tailored to Gikuriro's stated target population: namely, households that contained malnourished children, or pregnant and lactating mothers. A core challenge of the benchmarking endeavor is the need to use a measure of eligibility in a manner that can be defined identically across arms.[40] As a result, we established a set of 'hard' eligibility criteria on the basis of which beneficiaries would be selected and the survey would be stratified. Households meeting these criteria would be identified by the survey firm, Innovations for Poverty Action (IPA), prior to sampling for the baseline study, to establish a comparable population of eligible households in all arms—including control—of the study.

CRS and USAID agreed that the following criteria represent the target population for Gikuriro:

- Criteria 1. All households in a village with a malnourished child (defined by a threshold value of weight/age) were enrolled.
  - Weight/age is used because it is believed that this data is more consistently available than data on middle-upper arm circumference (MUAC) and height/age, and because it is used by CHWs as a basis for referring children to their local Health Centers.
  - The threshold weight/age value for inclusion was determined using the Rwandan Ministry of Health standards for malnutrition. The data used to identify eligibles was based on the Community Health Worker data from Growth Monitoring and Promotion visits.
- Criteria 2. All households in Ubudehe 1 or 2 with children under the age of 5 (Ubudehe is the Rwandan government household-level poverty classification, with 1 being the poorest, 3 being non-poor, and rural areas containing very few of the wealthiest Ubudehe 4 households).
- Criteria 3. All households in Ubudehe 1 or 2 with a pregnant or lactating mother.

Both implementers agreed to attempt to treat all eligible households that were identified as meeting any of these criteria. CRS anticipated an average of 30 eligible households per village, and in

[40] We did not intend the scope of the benchmarking exercise to include the implementers' (potentially different) ability to cost-effectively identify this target population, so as to maintain the interpretation of impacts as being differential impacts on a consistently defined beneficiary group.

principle had established a rationing rule in case that number was exceeded. As will be described below, the number of households per village that could be identified by the survey firm as meeting these targets turned out to be substantially lower. We did not try to impose restrictions on how Gikuriro would target outside of the households identified by the survey firm to be eligible.

We asked IPA to identify the universe of households that they could locate who met these criteria, using three sources. First, CHW records from the national 'Growth Monitoring and Promotion' exercise, which is intended to provide monthly height and weight measurements for all children under two and annual measurements for all children under five; second, government (census) records of household *Ubudehe* classifications; and finally local health facility information, which provides an alternative data point on children's nutritional status.[41] Children were defined as malnourished if they had at least one measurement that met government thresholds for malnourishment definitions in the past year, and households were defined as eligible if they had any individual meeting the criteria above. In each village we recorded the number of households in each stratum and sampled up to eight eligibles and four ineligibles for inclusion in the study. Throughout this document we use the words 'eligible' and 'ingeligible' to refer to the classification made by the survey firm at baseline.

[41] In practice, most children attending local clinics are referred by a CHW and so are also recorded as malnourished in the Growth Monitoring process.

Appendix C. Attrition

Endline outcome measurement is subject to a number of distinct forms of attrition; we start our empirical analysis by considering each in turn. The most straightforward of these is standard household-level attrition, meaning that a household sampled into the baseline survey attrited from the endline survey. In Table A.2, we see that overall rates of attrition at the household level were low, around 3.3 percent in the control. We see the pattern typical in RCT studies where attrition is somewhat lower in the treatment groups (where both ongoing contact and a sense of reciprocity may keep individuals in the endline), but these differentials are small, from 0.89 percentage points in the GD 'small' arm to 1.7 percentage points in the GD 'large' arm; only the latter is significant, and only at the 10 percent level. Looking at the other covariates of attrition in column 2 we see that attriters and non-attriting households are similar. Hence we conclude that household-level attrition is unlikely to be a source of bias in the study.

When we turn to the analysis of individual-level outcomes in Columns (3)–(9) the picture is more complex because many of the primary and secondary outcomes are only measured for certain types of individuals (anthropometrics for children, birth outcomes only for those pregnant). We analyze four types of individual missingness that may occur. First, we compare the attrition of all household members from the roster in the household survey; both the rates and the differentials here are very similar to the household attrition problem suggesting that there has been little additional differential attrition of individuals. Next we examine the anthropometric panel, whereby all children under 6 at baseline who were given anthropometrics at the baseline should have been followed up with at endline. Here the absolute rates of attrition are a little more than double what they are for individuals overall, presumably because of the greater difficulty of finding and measuring children for this exercise. More concerningly, the decline in attrition in the treatment groups now becomes strongly significant, particularly for Gikuriro villages (which perhaps is evidence of the superior monitoring of malnourished children taking place in those villages). Given this significance, we follow our pre-analysis plan in also presenting results for the anthropometric impacts that are corrected by inverse propensity weights to correct for the observable determinants of selection, and we also present Lee Bounds for this attrition. Third, we examine whether individuals who should have been anemia tested in the followup were; here we see no evidence of differential attrition across

Table C.1: Balance on Primary Outcomes

	Gikuriro Village	GD Main Village	GD Large Village	Control Mean	Observations	R^2
A. Household outcomes						
Consumption†	0.053	0.047	−0.103	10.39	1751	0.05
	(0.116)	(0.122)	(0.130)			
	[1.00]	[1.00]	[1.00]			
Household dietary diversity score	−0.064	−0.071	−0.058	4.16	1751	0.10
	(0.137)	(0.139)	(0.172)			
	[1.00]	[1.00]	[1.00]			
Household non-land wealth†	0.027	−0.042	−0.288	12.94	1751	0.06
	(0.227)	(0.214)	(0.254)			
	[1.00]	[1.00]	[1.00]			
B. Individual outcomes						
Height-for-Age	0.046	0.080	0.215**	−1.93	2187	0.02
	(0.086)	(0.097)	(0.098)			
	[1.00]	[1.00]	[0.13]			
Weight-for-Age	0.023	0.041	0.187***	−1.06	2180	0.02
	(0.069)	(0.070)	(0.069)			
	[1.00]	[1.00]	[0.07]			
Mid-Upper Arm Circ	0.015	0.025	0.070	−0.72	1987	0.04
	(0.068)	(0.068)	(0.081)			
	[1.00]	[1.00]	[1.00]			

Notes: Columns present coefficients and standard errors from a regression of each baseline outcome on treatment indicators, with fixed effects for blocks. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets. Variables marked with a † are in inverse hyperbolic sines.

arms. Finally, we examine the likelihood that a new household member appears (typically due to births subsequent to baseline), and find no significant differences. Overall, then, differential selection across treatment arms is not a major problem for study outcomes other than anthropometrics. We return to the issue of unequal attrition in anthropometrics in the following section.

Table C.2: Balance on Secondary Outcomes

	Gikuriro Village	GD Main Village	GD Large Village	Control Mean	Observations	R^2
A. Household outcomes						
Stock of borrowing†	−0.459	−0.007	−0.262	5.96	1751	0.04
	(0.363)	(0.409)	(0.408)			
	[0.89]	[1.00]	[1.00]			
Stock of saving†	−0.157	−0.665*	−0.269	5.18	1751	0.02
	(0.378)	(0.364)	(0.421)			
	[1.00]	[0.89]	[1.00]			
Health knowledge index	−0.590	−0.119	−0.225	0.19	1751	0.03
	(0.366)	(0.412)	(0.520)			
	[0.89]	[1.00]	[1.00]			
Sanitation practices index	0.285*	−0.105	−0.069	−0.23	1751	0.04
	(0.169)	(0.190)	(0.210)			
	[0.89]	[1.00]	[1.00]			
Productive assets†	0.281**	0.195	0.231*	11.41	1751	0.12
	(0.125)	(0.132)	(0.122)			
	[0.89]	[0.89]	[0.89]			
Consumption assets†	0.158	−0.034	0.426	8.71	1751	0.08
	(0.290)	(0.316)	(0.300)			
	[1.00]	[1.00]	[0.89]			
House value†	−0.042	−0.012	−0.067	13.59	1751	0.09
	(0.059)	(0.074)	(0.066)			
	[1.00]	[1.00]	[1.00]			
Housing quality index	0.018	−0.195	−0.014	0.02	1751	0.04
	(0.112)	(0.132)	(0.198)			
	[1.00]	[0.89]	[1.00]			
B. Individual outcomes						
Pregnancy	−0.018	−0.031	−0.021	0.28	2358	0.03
	(0.025)	(0.022)	(0.024)			
	[1.00]	[1.00]	[1.00]			
Live Birth	−0.017	−0.007	0.085	0.81	645	0.10
	(0.050)	(0.049)	(0.061)			
	[1.00]	[1.00]	[1.00]			
Birth in Facility	0.011	−0.056	−0.024	0.93	544	0.11
	(0.038)	(0.043)	(0.044)			
	[1.00]	[1.00]	[1.00]			
Any Vaccinations in past year	0.009	−0.006	0.001	0.93	1349	0.01
	(0.019)	(0.021)	(0.030)			
	[1.00]	[1.00]	[1.00]			
Completed Vaccinations	−0.015	−0.015	0.017	0.72	1347	0.02
	(0.037)	(0.045)	(0.042)			
	[1.00]	[1.00]	[1.00]			
Disease Burden	0.030	0.004	0.007	0.42	1146	0.02
	(0.040)	(0.032)	(0.043)			
	[1.00]	[1.00]	[1.00]			

Notes: See prior table. Indexes are unweighted sums of z-scores of their underlying components. Individual secondary outcomes all measured as rates within respective populations.

Table C.3: Balance on Household Covariates

	Gikuriro Village	GD Main Village	GD Large Village	Control Mean	Observations	R^2
Female Headed	0.036	0.043*	−0.018	0.16	1751	0.06
	(0.025)	(0.026)	(0.029)			
	[1.00]	[0.84]	[1.00]			
Agricultural	0.017	−0.027	0.002	0.85	1751	0.04
	(0.028)	(0.029)	(0.035)			
	[1.00]	[1.00]	[1.00]			
Wage Worker	−0.002	−0.063**	−0.084**	0.25	1751	0.04
	(0.029)	(0.031)	(0.035)			
	[1.00]	[0.63]	[0.46]			
Microenterprise	−0.015	0.008	−0.024	0.13	1751	0.02
	(0.025)	(0.024)	(0.023)			
	[1.00]	[1.00]	[1.00]			
Savings Group	−0.013	−0.022	0.026	0.25	1751	0.02
	(0.038)	(0.039)	(0.044)			
	[1.00]	[1.00]	[1.00]			
Village Eligibility Ratio	−0.015	0.037	0.017	0.16	1751	0.50
	(0.025)	(0.029)	(0.033)			
	[1.00]	[1.00]	[1.00]			
Age of Head	2.186**	2.868**	1.415	34.16	1751	0.07
	(1.047)	(1.200)	(1.487)			
	[0.63]	[0.46]	[1.00]			
Schooling of Head	−0.006	−0.002	−0.005	0.00	1751	0.02
	(0.005)	(0.006)	(0.004)			
	[1.00]	[1.00]	[1.00]			
Dependency Ratio	0.008	−0.007	0.003	0.59	1751	0.04
	(0.012)	(0.012)	(0.016)			
	[1.00]	[1.00]	[1.00]			
Household Size	−0.082	−0.054	−0.183	5.18	1751	0.02
	(0.134)	(0.151)	(0.163)			
	[1.00]	[1.00]	[1.00]			
Poorest Category	−0.040	−0.002	−0.068*	0.22	1751	0.05
	(0.033)	(0.045)	(0.039)			
	[1.00]	[1.00]	[0.84]			
Next Poorest Category	0.067*	0.056	0.061	0.50	1751	0.12
	(0.040)	(0.046)	(0.051)			
	[0.84]	[1.00]	[1.00]			

Notes: Columns present coefficients and standard errors from a regression of baseline covariates on treatment indicators, with fixed effects for blocks. Asterices denote significance at the 10, 5, and 1 percent levels, and are based on clustered standard errors, in parentheses. Anderson (2008) sharpened q-values presented in brackets.

Appendix D. Study Outcomes

Primary Outcomes. The study focuses on five dimensions. Here we briefly summarize each; details of the construction of these outcomes are included in Appendix A.

1. Household monthly consumption per capita (inverse hyperbolic sine—henceforth IHS—to deal with skewness).

2. Household Dietary Diversity, measured using hte WHO standard Household Dietary Diversity Score.

3. Anemia: measured with a biomarker test following DHS protocols at endline only.

4. Child growth and development: measured using in height-for-age, weight-for-age and Mid Upper Arm Circumference at baseline and endline for children under the age of 6 in eligible households.

5. Value of household non-land net wealth. This outcome is the sum of productive and consumption assets; the value of the household's dwelling, if owned; and the value of the stock of net savings, less the stock of debt (IHS).

Secondary Outcomes. Three types of outcomes are selected to be secondary: proximate outcomes of one or both interventions that do not have an intrinsic welfare interpretation (such as borrowing and saving stocks); outcomes that have welfare weight but are not within the causal chain of both programs (such as investments in health-seeking behavior, which Gikuriro seeks to impact, or housing quality, which has been identified as a dimension of benefit in prior evaluations of GiveDirectly (Haushofer and Shapiro, 2016)); or outcomes of common interest on which power is limited (such as disease burden and mortality).

1. Stock of borrowing and stock of savings (IHS).

2. Birth outcomes: the likelihood of pregnancy and likelihood of live birth within 12 months prior to endline.

3. Health knowledge and sanitation practices.

4. Disease burden and mortality. Mortality is measured as the likelihood that an individual member of the household from baseline has died prior to endline. Disease burden is measured as the prevalence of fever, fever with diarrhea or vomiting, or coughing with blood at endline,

5. Health-seeking behavior/preventative care. We focus on the share of pregnancies resulting in births in medical facilities, the share of children under two years of age with at least one vaccination in the prior year, and the share of children under two years of age with a complete dose of vaccines.

6. Household productive assets (IHS).

7. Housing quality. Two measures are used: the self-reported replacement cost of the current dwelling (irrespective of ownership status, IHS), and an index of housing construction quality, constructed from measures of wall and roof materials and from the number of rooms in the dwelling.

The inverse hyperbolic sine is commonly used in analysis of outcomes such as consumption, savings, and asset values that tend to be highly right-skewed and also to contain zeros. The IHS transformation preserves the interpretation of a log (meaning that impacts can be interpreted as percent changes) but does not drop zeros. Only outcomes that we expected to be skewed were pre-registered to be analyzed using IHS. All non-binary outcomes are also Winsorized at the 1 percent and 99 percent level (values above the 99th percentile are overwritten with the value at the 99th percentile to reduce skewness and increase statistical power). Because we restrict the analysis in this paper to the pre-specified primary and secondary outcomes only, we do not correct the results for multiple inference (Anderson, 2008).

Appendix E. Heterogeneous impacts of cash and kind

In this Appendix, we assess the extent of heterogeneity in ITT effects across observable subpopulations. Our pre-analysis plan highlighted two forms of heterogeneity that we anticipated would be important at the design phase; namely how baseline malnutrition and child age may moderate the impact of nutritional interventions. Given that we have children who start the study outside of the first 1,000 days (those 2–5 years old at baseline), we might expect that the impact of the program on these more fully developed children would be smaller. Similarly, we might expect that both of these interventions would be most effective for children who began the intervention most malnourished.

Our analysis generally reveals a lack of heterogeneity, in that impacts are not larger for children most malnourished at baseline (Table E.1), or for children exposed to the treatments at younger ages (Table E.2). Figures E.1a and E.1b provide fan regressions of impacts by child age and there is a suggestion that children exposed to large cash transfers in utero realize the largest benefits.[42] In general, however, our results are not suggestive of strong age- or malnutrition-driven heterogeneity of these interventions.[43]

[42] This pattern is similar to the medium-term results in Baird et al. (2016), who find unconditional transfers in Malawi to have the largest effect on children exposed in utero.

[43] The longer-term literature has typically found impacts of large cash transfer programs on HAZ in the range of 0.2–0.45 standard deviations (Aguero et al., 2006; Barham et al., 2014); comparison to the broader literature suggests that these impacts may grow over time.

Table E.1: Heterogeneity by Baseline Malnutrition

	(1) Height-for-Age	(2) Weight-for-Age	(3) Mid-Upper Arm Circ
Baseline outcome x Gikuriro	-0.0416	-0.0349	0.0852
	(0.0444)	(0.0619)	(0.0564)
Baseline outcome x GD Main	-0.0247	-0.0654	0.0776
	(0.0457)	(0.0445)	(0.0653)
Baseline outcome x GD Large	0.0220	0.00599	0.0804
	(0.0433)	(0.0461)	(0.0603)
Gikuriro	0.0434	0.0323	0.0253
	(0.0428)	(0.0362)	(0.0557)
GD Main	-0.0252	0.00182	-0.00498
	(0.0398)	(0.0357)	(0.0647)
GD Large	0.0940^{*}	0.0641	0.135^{*}
	(0.0517)	(0.0392)	(0.0795)
Baseline Outcome	0.768^{***}	0.748^{***}	0.600^{***}
	(0.0336)	(0.0355)	(0.0425)
Observations	2125	2104	1629
Mean of DV	-2.031	-1.043	-0.572
R squared	0.696	0.673	0.507

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Regressions with both baseline and endline outcome measurement are ANCOVA with lagged dependent variables as controls, run on the panel sample. Regressions include fixed effects for the randomization blocks, and are weighted to be representative of all households in study villages. Anthropometric outcomes are demeaned prior to interaction so that the uninteracted treatment terms provide impact at average level of baseline anthro measure..

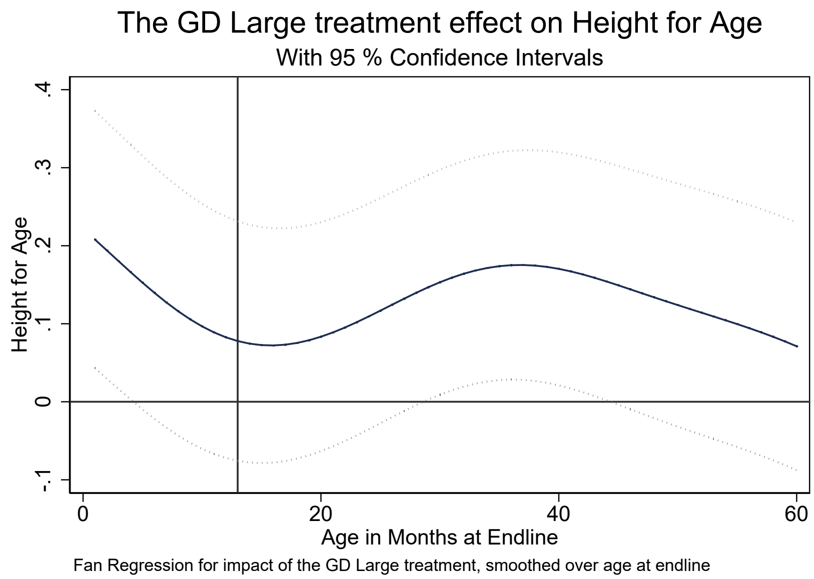
Table E.2: Heterogeneity by Baseline Age

	First Thousand Days (1) Height-for-Age	(2) Weight-for-Age	(3) Mid-Upper Arm Circ	Newborn (4) Height-for-Age	(5) Weight-for-Age	(6) Mid-Upper Arm Circ
Indicator x Gikuriro	-0.00731	-0.0206	0.115	0.599	0.251	0.282
	(0.138)	(0.113)	(0.109)	(0.645)	(0.505)	(0.491)
Indicator x GD Main	-0.300**	-0.152	0.159	0.382	0.594	0.666
	(0.138)	(0.115)	(0.104)	(0.522)	(0.495)	(0.506)
Indicator x GD Large	-0.115	-0.0159	0.160	0.407	0.729	0.304
	(0.139)	(0.122)	(0.144)	(0.396)	(0.469)	(0.281)
Gikuriro	0.0105	0.0114	-0.0813	0.00489	0.00325	-0.0263
	(0.106)	(0.0801)	(0.0822)	(0.0833)	(0.0629)	(0.0690)
GD Main	0.115	0.0779	-0.108	-0.0171	0.00905	-0.0346
	(0.119)	(0.0829)	(0.0905)	(0.0988)	(0.0678)	(0.0723)
GD Large	0.246**	0.191**	0.0773	0.196**	0.185***	0.159**
	(0.105)	(0.0817)	(0.111)	(0.0848)	(0.0668)	(0.0783)
Indicator	0.141	0.123	-0.0101	-0.0102	0.0247	0.177
	(0.148)	(0.117)	(0.142)	(0.254)	(0.295)	(0.270)
Observations	2360	2347	2020	2360	2347	2020
Mean of DV	-2.031	-1.043	-0.572	-2.031	-1.043	-0.572
R squared	0.0722	0.0356	0.0726	0.0699	0.0358	0.0740

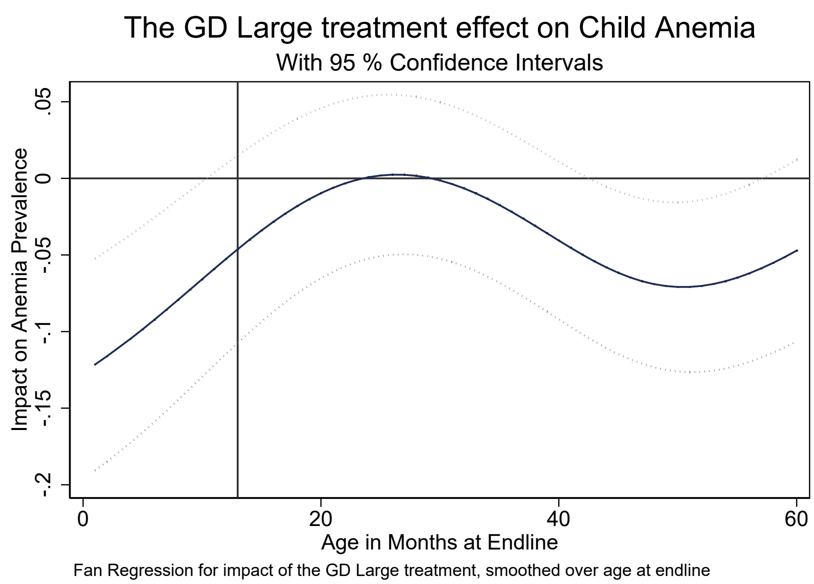
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: First three columns present an interaction with an indictor for a child in the 'First Thousand Days' (<33 months at endline) and the last three columns present interactions with an indicator for 'Newborn' (< 13 months at endline). Regressions are endline cross-sections, run on the panel sample, and do not include the lagged outcome variable so as to be able to consider children who are newborns in R2. Regressions include fixed effects for the randomization blocks, and are weighted to be representative of eligible households in study villages.

Figure E.1: Fan Regression Impacts by Age



(a) Impacts on Height-for-Age



(b) Impacts on Child Anemia

Appendix F. Selection of Control Variables.

In our pre-analysis plan, we state that control variables for the primary specification "will be selected on the basis of their ability to predict the primary outcomes". In doing so, we seek to build on recent developments that balance the challenge of using baseline data to select variables that will reduce residual variance with the danger that researcher freedom in the selection of control variables can lead to p-hacking, in which right-hand-side variables are selected specifically on the basis of the statistical significance of the coefficient of interest (Card and Krueger, 1995; Casey et al., 2012), thereby invalidating inference.

To balance these concerns, we follow the *post-double-selection* approach set forth in Belloni et al. (2014b). Those authors advocate a two-step procedure in which, first, Lasso is used to automate the selection of control variables, and second, the post-Lasso estimator (Belloni et al., 2012) is used to estimate the coefficients of primary interest in in the ITT, effectively using Lasso as a model selection device but *not* imposing the shrunken coefficients that results from the Lasso estimates directly. Belloni et al. (2014b) demonstrate that this approach not only reduces bias in estimated treatment effects better than alternative approaches—less a concern given the successful randomization in our experiment—but that it may improve power while retaining uniformly valid inference.

In the first stage, model selection is undertaken by retaining control variables from the union of those chosen either as predictive of the treatment assignment or of the outcome. This model selection stage can be undertaken after residualizing to account for a set of control variables that the authors have a priori determined belonw in the model, as in Belloni et al. (2014a); in our case, we retain block fixed effects, lagged values of the outcome, and lagged values of (the inverse hyperbolic sine of) household wealth in all specifications, per our pre-analysis plan. We modify the heteroskedasticity-robust Lasso estimator of Belloni et al. (2012) to incorporate sampling weights consistent with our design, using the Lasso penalty is chosen as a function of the sample size and the number of potential covariates, as in Belloni et al. (2014a).

Resulting covariates selected for each of the primary and secondary outcomes, at household and individual level, are presented in Tables F.1 and F.2, respectively.

Table F.1: Covariates selected in Belloni et al. (2014) post-double-lasso selection procedure for household outcomes

Outcome	Control set
consumption_asinh	Baseline value of consumption_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
Household dietary diversity score	Baseline value of dietarydiversity, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	Lsavingsstock_asinh3
	Lconsumpti_x_Ldietarydi
	Lconsumpti_x_Lproductiv
	Ldietarydi_x_Lassetscon
wealth_asinh	Baseline value of wealth_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	L.Own dwelling
borrowingstock_asinh	Baseline value of borrowingstock_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
savingsstock_asinh	Baseline value of savingsstock_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	Lconsumpti_x_Lproductiv
	Lconsumpti_x_Lassetscon
Health Knowledge Index	Baseline value of health_knowledge, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
Sanitation Practices Index	Baseline value of sanitation_practices, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	Lproductiv_x_Lassetscon
productiveassets_asinh	Baseline value of productiveassets_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	Lconsumpti_x_Lassetscon
assetsconsumption_asinh	Baseline value of assetsconsumption_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	L.Number of rooms
	L.Durables expenditure (12-month recall)
	Ldietarydi_x_Lassetscon
	Lproductiv_x_Lassetscon
selfcostdwell_asinh	Baseline value of selfcostdwell_asinh, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	L.Number of rooms
	L.Durables expenditure (12-month recall)
Housing Quality Index	Baseline value of housing_quality, present in both rounds
	L.Lhh_wealth_asinh
	L.Fraction of village defined eligible by IPA
	L.Number of rooms

Note: block fixed effects and lag of the relevant outcome included in all specifications. Specifications that include both eligible and ineligible households include an indicator for eligibility status.

Table F.2: Covariates selected in Belloni et al. (2014) post-double-lasso selection procedure for individual outcomes

Outcome	Sample	Control set
haz06, Winsorized fraction .005, high only	Under 5s	L.haz06, Winsorized fraction .005, high only
		female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		L.Food expenditure (weekly recall)
		L.Food consumption-value own production (weekly recall)
		L.waz06, Winsorized fraction .005, high only
		Lconsumpti_x_Lselfcostd
waz06, Winsorized fraction .005, high only	Under 5s	L.waz06, Winsorized fraction .005, high only
		female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		L.Food expenditure (weekly recall)
		L.Food consumption-value own production (weekly recall)
		Lconsumpti_x_Lproductiv
muacz, Winsorized fraction .01	Under 5s	L.muacz, Winsorized fraction .01
		female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		L.waz06, Winsorized fraction .005, high only
		Lconsumpti_x_Lproductiv
anemia_dummy	Under 5s	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
anemia_dummy	Pregnant/lactating women	agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
mortality	All	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
Was this women pregnant at any point in the past 12 months	Pregnant/lactating women	agemonths
		agemonths_sq
		agemonths_cu

Continued on next page

Table F.2 (continued)

Outcome	Sample	Control set
		L.Lhh_wealth_asinh
		L.Lwealth_asinh
Did pregnancy conclude in live birth	Pregnant/lactating women	agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		L.Food expenditure (weekly recall)
		L.Food consumption-value own production (weekly recall)
		Lconsumpti_x_Lwealth_as
facility_birth	Pregnant/lactating women	agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
anthro_vacc_year	Under 3s	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		Lconsumpti_x_Lproductiv
anthro_vacc_complete	Under 3s	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
Any fever, diarrhea, or coughing blood at individual/round level	Under 5s	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh
		L.Food consumption-value own production (weekly recall)
Individual reported with diarrhea/vomiting/fever now	Under 5s	female
		agemonths
		agemonths_sq
		agemonths_cu
		L.Lhh_wealth_asinh

Note: block fixed effects and lag of the relevant outcome included in all specifications. Specifications that include both eligible and ineligible households include an indicator for eligibility status.