RESEARCH ARTICLE SUMMARY

DEVELOPMENT ECONOMICS

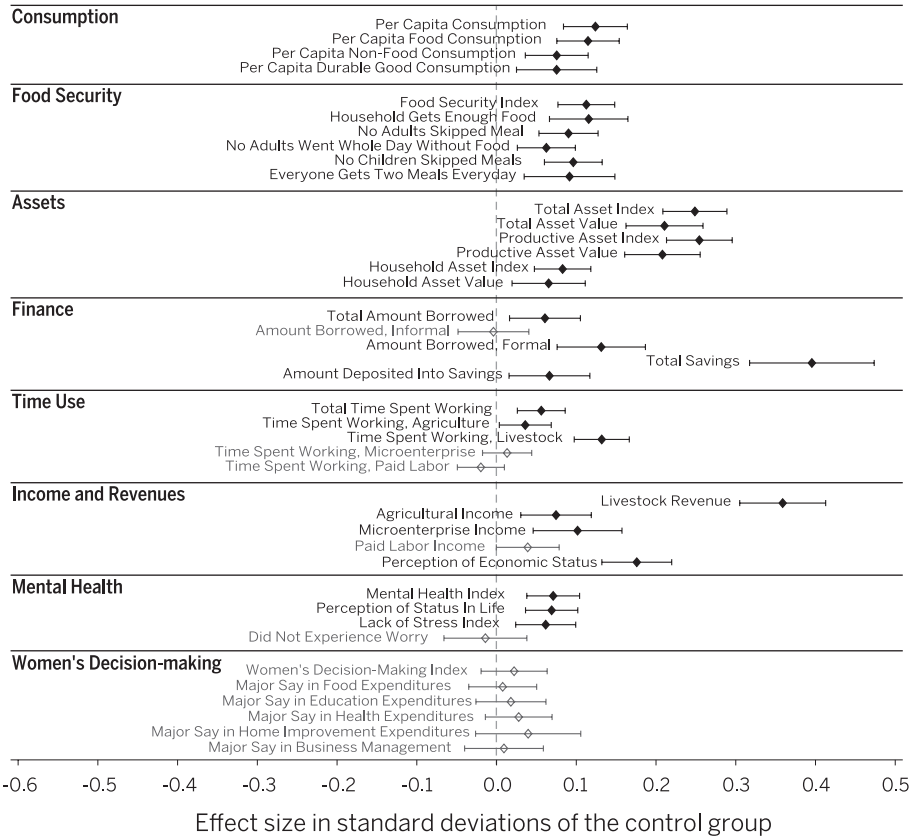
A multifaceted program causes lasting progress for the very poor: Evidence from six countries

Abhijit Banerjee, Esther Duflo, Nathanael Goldberg, Dean Karlan,* Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, Christopher Udry

INTRODUCTION: Working in six countries with an international consortium, we investigate whether a multifaceted Graduation program can help the extreme poor establish sustainable self-employment activities and generate lasting improvements in their well-being. The program targets the poorest members in a village and provides a productive asset grant, training and support, life skills coaching, temporary cash consumption support, and typically access to savings accounts and health information or services. In each country, the program was adjusted to suit different contexts and cultures, while staying true to the same overall principles. This multipronged approach is relatively expensive, but the theory of change is that the combination of these activities is necessary and sufficient to obtain a persistent impact. We do not test whether each of the program dimensions is individually necessary.

Pooled average intent-to-treat effects, endline 2 at a glance

This figure summarizes the average treatment effects in each country for the 10 primary outcomes. All treatment effects are presented as standardized *z*-score indices and 95% confidence intervals.



Instead, we examine the "sufficiency" claim: A year after the conclusion of the program, and 3 years after the asset transfer, are program participants earning more income and achieving stable improvements in their well-being?

RATIONALE: We conducted six randomized trials in Ethiopia, Ghana, Honduras, India, Pakistan, and Peru with a total of 10,495 participants. In each site, our implementing partners selected eligible villages based on being in geographies associated with extreme poverty, and then identified the poorest of the poor in these villages through a participatory wealth-ranking process. About half the eligible participants were assigned to treatment, and half to control. In three of the sites, to measure within village spillovers, we also randomized half of villages to treatment and half to control. We conducted a baseline survey on all eligible participants, as well as an endline at the end of the intervention (typically 24 months after the start of the intervention) and a second endline 1 year after the first endline. We measure impacts on consumption, food security, productive and household assets, financial inclusion, time use, income and revenues, physical health, mental health, political involvement, and women's empowerment.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1260799

RESULTS: At the end of the intervention, we found statistically significant impacts on all 10 key outcomes or indices. One year after the end of the intervention, 36 months after the productive asset transfer, 8 out of 10 indices still showed statistically significant gains, and there was very little or no decline in the impact of the program on the key variables (consumption, household assets, and food security). Income and revenues were significantly higher in the treatment group in every country. Household consumption was significantly higher in every country except one (Honduras). In most countries, the (discounted) extra earnings exceeded the program cost.

CONCLUSION: The Graduation program's primary goal, to substantially increase consumption of the very poor, is achieved by the conclusion of the program and maintained 1 year later. The estimated benefits are higher than the costs in five out of six sites. Although more can be learned about how to optimize the design and implementation of the program, we establish that a multifaceted approach to increasing income and well-being for the ultra-poor is sustainable and cost-effective. ■

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: dean.karlan@yale.edu
Cite this article as A. Banerjee *et al.*, *Science* **348**, 1260799 (2015). DOI: 10.1126/science.1260799

RESEARCH ARTICLE

DEVELOPMENT ECONOMICS

A multifaceted program causes lasting progress for the very poor: Evidence from six countries

Abhijit Banerjee,[1,2,3,4] Esther Duflo,[1,2,3,4] Nathanael Goldberg,[5] Dean Karlan,[2,3,4,5,6]* Robert Osei,[7] William Parienté,[4,8] Jeremy Shapiro,[9] Bram Thuysbaert,[5,10] Christopher Udry[2,3,4,6]

We present results from six randomized control trials of an integrated approach to improve livelihoods among the very poor. The approach combines the transfer of a productive asset with consumption support, training, and coaching plus savings encouragement and health education and/or services. Results from the implementation of the same basic program, adapted to a wide variety of geographic and institutional contexts and with multiple implementing partners, show statistically significant cost-effective impacts on consumption (fueled mostly by increases in self-employment income) and psychosocial status of the targeted households. The impact on the poor households lasted at least a year after all implementation ended. It is possible to make sustainable improvements in the economic status of the poor with a relatively short-term intervention.

More than one-fifth of the world's population lives on less than purchasing power parity (PPP) US$1.25 a day, and there is an emerging international consensus that this share should (and can) be driven close to zero by 2030 (*1*, *2*). Reaching this objective will require enabling the poorest families, who are often the most marginalized within their villages, to shift from insecure and fragile sources of income to more sustainable income-generating activities. One possible avenue, popular with both development organizations and governments, is to promote self-employment activities (such as cow rearing or petty trading). Past efforts to reduce poverty by encouraging these types of activities among the poor, however, have often been plagued by implementation problems and been deemed failures (*3*). For example, India's Integrated Rural Development Program (IRDP) is believed to have been both poorly targeted and ineffective (*4*, *5*). However, in recent years, several large nongovernmental organizations (prominent international northern NGOs such as Oxfam, World Vision, and Heifer, as well as many local NGOs) have gone back to this "livelihood" approach. This past experience raises the question: Is it actually possible to reliably improve the livelihoods of the poorest households by giving them access to self-employment activities, or is this entire approach flawed? In particular, is it possible to come up with a model for doing so that can be implemented by a wide variety of organizations and works in a wide range of geographic, institutional, and cultural contexts?

We present results from randomized control trials (RCTs) in six countries of a particular approach to foster self-employment activities among the very poor. Originally designed and implemented by BRAC, a large Bangladeshi NGO that runs several country-wide programs, the "Graduation" program provides a holistic set of services, including the grant of a productive asset, to the poorest households in a village (referred to by BRAC as the "ultra-poor"). The beneficiaries are identified through a participatory process in a village meeting, followed by a verification visit by the organization's staff. Selected beneficiaries are then given a productive asset that they choose from a list, training and support for the asset they have chosen, as well as general life skills coaching, weekly consumption support for some fixed period, and typically access to savings accounts and health information or services. These different activities (plus regular interactions with the households over the course of a year) are designed to complement each other in helping households to start a productive self-employment activity. The idea is to provide a "big push," over a limited period of time, with the hope of unlocking a poverty trap. The program costs per household average 100% (range from 62 to 145%) of baseline household consumption. Although the program may initially be relatively expensive (compared to just providing training, coaching or a cash transfer), the thinking behind the program is that the combination of these activities is necessary and sufficient to obtain a persistent impact on a large fraction of the beneficiaries.

We address the "sufficiency" claim: Is the Graduation approach effective and cost-effective, and can it be implemented at scale and in different contexts and cultures? Whether all the ingredients of the program are individually necessary is not tackled here and will need to be dealt with in future work.

A key feature of the BRAC approach is that, while comprehensive, it is well codified, scalable, and replicable. BRAC has already implemented the program at scale in Bangladesh. As of 2011, BRAC had reached close to 400,000 households, and a further 250,000 were scheduled to be reached between 2012 and 2016 (*6*). It has now also been replicated in about 20 countries, including the six countries that are studied here. A high-quality RCT, conducted independently but simultaneously with this study, has shown the BRAC program in Bangladesh to be very effective (*6*). Two years after graduation, households have expanded their self-employment activities, diversified out of agriculture and livestock, reduced casual labor, and increased consumption. Previous nonrandomized studies of the BRAC program (*7*–*9*) found similar impacts.

Between 2007 and 2014, we conducted a multisite RCT of the Graduation program. The sites were chosen as part of an effort led by the Ford Foundation and Consultative Group to Assist the Poor (CGAP), referred to here as the Graduation Program Consortium. The programs were implemented by six different organizations in six countries (Ethiopia, Ghana, Honduras, India, Pakistan, and Peru), but overall planning on the programs and evaluation were coordinated from the onset (*10*). Treatment was randomly assigned among eligible households. Data were collected at baseline and immediately after the programs ended, 2 years later ("endline 1"), and again 1 year after the programs ended, i.e., about 3 years after the beginning of the programs ("endline 2"). We report pooled results from all the sites (21,063 adults in 10,495 households), as well as site-by-site results.

The main contribution of this study is the evaluation of the cost-effectiveness of the same potentially important intervention across a diverse set of contexts. The sites span three continents, and different cultures, market access and structures, religions, subsistence activities, and overlap with government safety net programs. This diversity should give us a high level of confidence in the robustness of the impact to variations in both the context and implementation agency. The core components of the program are similar in substance and magnitude, although the program design includes adjustments as are necessary for local contexts. For example, country-specific market analysis was conducted to determine viable livelihoods to promote, rather than simply promoting the same livelihood in every context. In addition, because the study was conceived from

[1]Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. [2]Center for Economic and Policy Research (CEPR), Washington, DC, USA. [3]National Bureau of Economic Research (NBER), Cambridge, MA, USA. [4]Jameel Poverty Action Lab (J-PAL), MIT, Cambridge, MA, USA. [5]Innovations for Poverty Action (IPA), New Haven, CT, USA. [6]Yale University, New Haven, CT, USA. [7]University of Ghana–Legon, Accra, Ghana. [8]Institut de Recherches Economiques et Sociales (IRES), Université Catholique de Louvain, Louvain-La-Neuve, Belgium. [9]Princeton University, Princeton, NJ, USA. [10]Ghent University, Ghent, Belgium.
*Corresponding author. E-mail: dean.karlan@yale.edu

the onset as one multisite study, variables were collected in a comparable manner on a broad array of outcomes. Finally, households were surveyed over 3 years, including 1 year after the end of the program, which directly speaks to the sustainability of the changes we observe.

The program: Commonalities and variations

The basic approach of the program is to combine six different activities designed to complement each other to help households start, and continue with, a self-employment activity. The core of the program is a productive asset transfer, but the premise of the program is that the support has to be sufficiently broad and long-lasting to ensure that households continue to benefit from that asset into the future.

Following identification of the beneficiary households through a participatory process in the village, the six activities are:

1. Productive asset transfer: a one-time transfer of a productive asset
2. Consumption support: a regular transfer of food or cash for a few months to about a year (*11*)
3. Technical skills training on managing the particular productive assets
4. High-frequency home visits
5. Savings: access to a savings account and in some instances a deposit collection service and/or mandatory savings
6. Some health education, basic health services, and/or life-skills training

The Graduation Program Consortium organized global learning events at which staff from each of the sites, along with researchers, gathered to discuss site-specific design considerations. The Consortium also hosted a dedicated website to foster ongoing knowledge exchange between sites and a wider community of practice. There were five global learning events between 2008 and 2014, plus several regional workshops. The first two global meetings featured exposure visits to the BRAC program in Bangladesh and the Bandhan program in India. Each partner thus participated in at least two field visits, with some additional exchange visits arranged on an ad hoc basis (e.g., the Ghana team visited the Ethiopia site as they designed their program).

We now detail the core components of the program. We first discuss the commonalities across all sites, and then discuss the important variations across sites. Table 1 has a detailed description of the program features in each site.

Targeting

The Graduation program is intended to serve the poorest of the poor within villages. The targeting process starts with selection of a poor region based on national survey data, and a list of villages within the target area (often selected in consultation with program staff). At most program sites, ultrapoor households are then identified using a Participatory Wealth Ranking (PWR) during which villagers create an economic ranking of all village households. In Indonesia, Alatas *et al.* (*12*) find that a PWR used to identify recipients of a government program successfully identified the poorer households. The households selected for the Graduation program through the PWR are then visited by field officers from the implementing organizations to verify their poverty status with an asset checklist [often the Progress out of Poverty (PPI) scorecard (*13*)]. Of the selected households, 48% have daily per capita consumption below PPP US$1.25, compared to 19% of the population at-large in these countries (table S1a).

A fraction of households in the resulting list are then randomly assigned to receive the program and are invited to participate. In all sites but India, all intended beneficiaries enrolled. We provide more discussion of take-up in the India program below.

Productive-asset transfer

The asset transfer is the core component of the program and also one of its largest costs. Each household chose, in consultation with the field officer, one of the assets (or asset bundle options) in a list proposed by the implementing organization (often, this list was created after hiring local experts to analyze markets and the viability of livelihood options). Common choices included raising livestock (sheep, goats, chicken, cattle, etc.) and petty trade, and are detailed in Table 1. The value of assets varied between sites, ranging from PPP US$437 to PPP US$1228 per household. The differences in transfer costs partially reflect the differences in local livestock prices: All but one site (Peru) transferred productive assets worth between four to eight goats at local prices (see Table 1 for exact figures). Furthermore, although the asset type differed across countries, the principle in choosing the asset was consistent. In four of the six sites, the asset transferred was the most or the second most commonly held asset at baseline. In Peru and Ethiopia, the most commonly transferred assets were guinea pigs, and sheep and goats, respectively, because they were believed to be more profitable than the most commonly held assets. Different assets generated quite different cash flow patterns: Some produced immediate revenue (e.g., petty trade), whereas others (like cows) produced far more delayed and lumpy revenue flows.

The asset transfer generally happened between 0 and 15 months, largely depending on the site, after the identification of the beneficiaries and the baseline survey. In Pakistan, where the intervention was run by several organizations, it took several months, and in some cases a year or more, to complete all rounds of asset transfers. Honduras also had delays in starting the program. In Ethiopia, the transfers were spread out over 6 months.

Consumption support

Consumption support—generally a cash stipend—was distributed typically weekly or monthly. The purpose of the consumption stipend is both to immediately improve and stabilize consumption, and to reduce incentives to sell (or eat up) the productive assets being distributed. The distribution of consumption support lasted between 4 and 13 months, depending on the site, and ranged from PPP US$24 to PPP US$72 per month (*14*). This variation partly reflects the fact that the PPP in each country is not based on the bundle of goods purchased by the poor: In all sites but Ethiopia (where the consumption support was part of an existing program), the transfer corresponds roughly to the monetary equivalent of between 2,402 and 5,142 calories per day (or roughly a kilogram of rice at local price) (*15*).

Consumption support was provided everywhere, but in two sites (Ethiopia and Peru), a form of consumption support already existed before the program started, so it was available for all (Ethiopia) or part (Peru) of the control group, as well. In Ethiopia, both treatment and control households received benefits from the Productive Safety Net Programme (PSNP), a food-for-work program for food-insecure households. For this reason, the program did not offer any additional consumption support to treatment households. In Peru, a conditional cash transfer program, Juntos, was active in 51 of the 86 project villages. Juntos provides PEN 200 (PPP US$143.33) every 2 months, on the condition that female heads of households meet the following conditions: obtain identity cards for their children, take children under 5 to health check-ups, and send children to school. In the non-Juntos villages, the treatment households received a "Juntos-like" consumption support: PEN 100 (PPP US$71.96) per month for 9 months, conditional on children attending school and receiving health check-ups. In our sample, 57% of control households report receiving support from Juntos during the baseline survey, whereas all the treatment households receive either Juntos or the replacement. Thus, Peru is an intermediate case between Ethiopia and the other sites.

Honduras implemented its consumption support by providing a one-time food transfer intended to cover the 6-month lean season.

Training

Before receiving their assets, households were provided with training on running a business and managing their chosen livelihood. For example, those selecting livestock received information on how to rear the livestock, including vaccinations, feed, and treatment of diseases.

High-frequency home visits

Households received regular training and coaching from a field officer throughout the 2-year program. The visits were intended to provide accountability (i.e., making sure that the households carry out the tasks necessary to maintain and grow their livelihood into a stable income-generating activity), as well as to be encouraging (e.g., helping households believe that they can have control of their lives and put themselves on a path out of extreme poverty) (*16*). During the home visits, field staff provided health education and financial capabilities coaching. In Peru, where traveling to the villages proved to be logistically challenging, visits happened only every 6 weeks, and in Pakistan, similar difficulties led the implementing NGOs to shift gradually to biweekly or monthly visits.

Table 1. Implementation summary.

	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
Location	Kilte Awlaelo district in Tigray	Northern and Upper East regions	Lempira department	Murshidabad, a district north of Kolkata	Sindh region	13 districts of the provinces of Canas and Acomayo, in the department of Cusco
Implementing NGO	Relief Society of Tigray (REST)	Presbyterian Agricultural Services (PAS) and Innovations for Poverty Action (IPA); program called Graduation from Ultra Poverty (GUP)	Proyecto MIRE, a partnership between PLAN International Honduras and Organización de Desarollo Empresarial Feminino Social (ODEF), a Honduran microfinance institution	Bandhan	Pakistan Poverty Alleviation Fund, Aga Khan Planning and Building Services Pakistan, Badin Rural Development Society, Indus Earth Trust, Sindh Agricultural and Forestry Workers Coordinating Organization	Asociación Arariwa; PLAN International Peru
Type of NGO	Local NGO	Local NGO	Local NGO; international NGO	Local MFI	Local NGOs	Local NGO; international NGO
Financial institution partner	Dedebit Credit and Savings Institution (DECSI), an Ethiopian microfinance institution (MFI)	Services provided by PAS	Services provided by ODEF	Services provided by Bandhan	None	Services provided by Arariwa
Eligibility requirements	Participant in food-for-work program, at least one member capable of work, no loans taken out by household	Exclusion criteria included: (i) ownership of >30 small ruminants or >50 fowl; (ii) member found to be alcoholic or drug addict; (iii) no strong, able-bodied adult; (iv) did not have a female member; (v) did not have a member between the ages of 18 and 65	Household must have (i) monthly income of less than or equal to HNL 600 (PPP US$67.82); (ii) lived in community for at least 3 years; (iii) not received a loan in last year; (iv) not a benificiary of cash or food transfer program (excluding government conditional cash transfer); (v) at least one child (under 18). Households must satisfy two of the following three criteria: (i) one manzana or less of cultivated land; (ii) children in the household work; (iii) not participating in development program	Household has able-bodied female member, households are not associated with any MFI and receive below a certain threshold of aid from the government. Meet three of the following criteria: primary source of income is informal labor or begging, land holdings are below 0.2 acres, household has no productive assets other than land, no able-bodied males in the household, school-age children work instead of attending school	Households meet at least three of the five selection criteria: no ownership of productive assets, no active male member in the households, income less than INR 25 (PPP US$1.74) per day, livelihood at risk, and no member working on salaried employment.	Household must (i) have head or spouse 60 years old or less; (ii) have at least one child under 18; (iii) have head not away from household for 6 months or more in a year; (iv) not borrow money from a formal financial institution; (v) not have a second home outside of the community; (vi) neither the head nor spouse has formal employment
Method of identifying participants	Chosen by local community's food security task force	Participatory Wealth Ranking at village level	Participatory Wealth Ranking at village level; followed by verification survey	Participatory Rural Appraisal, followed by NGO verificiation	Participatory Rural Appraisal, followed by NGO verification	Participatory Wealth Ranking at village level; followed by verification survey

Continued on next page

	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
Savings component	Households had bank accounts opened at DECSI, required to regularly deposit savings. In total, households required to save ETB 4724 (PPP US$1227.87) over the 2 years of the program; households were unable to withdraw funds until they reached this threshold	Half of GUP households received savings accounts (savings collected during weekly visits by field agents, households receive passbooks to log deposits)	Female heads of household are required to open savings accounts with ODEF and are randomly assigned to two savings treatments: (i) savings matching biannually equal to 50% of the avereage account balance, (ii) monthly direct savings transfers [both get savings incentives of up to HNL 800 (PPP US$90.42)]	Households required to save approximately INR 10 (PPP US$0.97) per week at weekly meeting with Bandhan staff	Households encouraged to save money at home, in boxes, or with Rotating Savings and Credit Associations (ROSCAs)	Participant households encouraged to join community savings groups; participants could also open passbooks with Banco de Nacion or deposit group savings with Arariwa Microfinance
Health component	None	Health and nutrition education. Beneficiaries were enrolled in the National Health Insurance Scheme.	Trainings in health, nutrition and hygiene	Health discussed at weekly meetings	Lady health visitors provided basic health services including checkups, health and hygiene training, and medicine. More difficult and serious cases were referred to the nearest doctor.	Health discussed at trainings (three modules: nutrition, healthy practices, prenatal health)
Asset transfer	June 2010–Jan 2011	July 2011–July 2012	August 2009–August 2010	February 2007–March 2008	December 2008–May 2010	February 2011–June 2011
Value of asset transfer	ETB 4724 (PPP US$1227.87)	GHS 300 (PPP US$451.38)	HNL 4750 (PPP US$536.89)	INR 4500 (PPP US$437.31)	PKR 15000 (PPP US$1043.33)	PEN 1200 (PPP US$853.97)
Value of asset transfer, in terms of local goat prices	7.98 goats	6.00 goats	4.75 goats	6.53 goats	3.75 goats	17.14 goats
Most common asset chosen	Sheep and goats (62%)	Goats and hens (44%)	Chickens (83%)	Goats (52%)	Goats (56%)	Guinea pigs (64%)
Second most common asset chosen	Oxen (24%)	Goats and maize inputs (27%)	Pigs (6%)	Cows (30%)	Shops (11%)	Hens (24%)
Third most common asset chosen	Bees (10%)	Shea nuts and hens (6%)	Fish (5%)	Nonfarm micro-enterprise inventory (11%)	Hens (10%)	Cattle (4%)
Consumption Support	Both treatment and control households received food support through food-for-work program through the duration of the program; 5 days of work (which can be completed once per month) earns 15 kg of wheat, 0.66 kg of chickpeas, and 0.4 liters of oil, worth approximately ETB 100 (PPP US$25.99)	Treatment households received weekly cash transfers of GHS 4 to 6 (PPP US$6.02 to 9.03) (amount dependent on household size) during lean season	Treatment houeholds received one-time food transfer worth HNL 1920 (PPP US$217.02), intended to cover the 6- month lean season	Treatment households received weekly cash transfers of INR 90 (PPP US$8.76) for 13 to 40 weeks, depending on the asset chosen (13 weeks for nonfarm enterprise, 30 weeks for goats, and 40 weeks for cows)	Treatment households received monthly cash transfers of PKR 1000 (PPP US$69.56) for the first year in the program	Treatment households that were not enrolled in a government conditional cash transfer program (Juntos) received monthly cash transfers of PEN 100 (PPP US$71.96). Treatment and control households enrolled in Juntos received bimonthly cash transfers of PEN 200 (PPP US$143.33)

Continued on next page

	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
Consumption support, measured in calories of cereals that could be purchased for the household if transfer were in cash	2,402	5,142	3,854	3,928	3,149	4,993
Baseline daily per capita consumption, measured in number of buyable calories of cereals	4,389	7,885	4,929	4,179	6,000	7,099
Consumption support provided to control households	Yes	No	No	No	No	Yes, in 51 of 86 project villages
Frequency of household visits from local NGO	Weekly, over 24 months	Weekly, over 24 months	Weekly, over 24 months	Weekly, over 18 months	Weekly, over 24 months, with gradual shift to biweekly or monthly in some cases	Every 6 weeks, over 24 months
Support provided from village assistance committees	No	Yes	Yes	Yes	No	No

Savings

Households were encouraged (and in some sites, required) to save in order to improve their ability to cope with shocks. This is one component that varied from site to site. Four sites (Ethiopia, Honduras, India, and Peru) partnered with microfinance institutions able to provide access to savings accounts. In Pakistan, households were encouraged to save through savings groups, and in Ghana, households received savings accounts. In India and Ghana, individuals were able to save at program meetings or with a visit by a field agent, but in the other four sites, households had to make deposits at the financial institution.

In Honduras, savings were further encouraged through financial incentives. Beneficiary households opened a savings account and were randomized into two groups: (i) savings matching semi-annually equal to 50% of the average account balance, or (ii) monthly direct savings transfers. Both groups received savings incentives equal to a maximum value of HNL 800 (PPP US$90.42). We do not analyze this experimental variation in this paper.

Ethiopia had a strong forced savings component. The government prohibited unconditional transfers to the poor. To satisfy this prohibition but still implement the program, the implementing partner, Relief Society of Tigray (REST) and the government agreed to allow the asset transfers to be described as "like" a loan, as recipients had to make deposits into a savings account in exchange for receiving the asset. Households were not able to withdraw their savings from the account until they saved an amount equal to ETB 4724 (PPP US$1228), the value of the asset transfer. However, once households achieved the required savings threshold, they had full access to their deposits and could withdraw from their accounts as they saw fit. Furthermore, if they failed to make the deposits, they did not forfeit their asset. Compliance with the deposits was very high, with only 15 households (out of 458) not fulfilling the commitment.

Health and other services

Finally, all sites but one (Ethiopia) included a health component such as health, nutrition, and hygiene training. Some sites also facilitated access to health care, either as direct services from community health workers, referring them to government or NGO health clinics, or by enrolling beneficiaries in national health insurance. Several of the sites organized support from village assistance committees comprising village leaders who helped advise the households, mediated problems, and connected beneficiaries with additional services.

Experimental methods

Experimental design

Of the six experiments, three are individual randomized trials with randomization at the household level within each village (India, Ethiopia, and Pakistan) and three are clustered randomized trials, with randomization at both the village and household level (Ghana, Honduras, and Peru). In the countries with clustered randomization, villages were randomly selected to be treatment or control villages, and then treatment households were randomly selected within the set of eligible households in treatment villages. The goal of this design was to be able to measure spillovers. For the main analysis in this paper, we ignore possible externalities and include all control households (within villages or across villages). In the Results section, we provide a discussion of whether any spillovers within the sample may bias our results. Randomization was carried out either remotely by the research team (using a computer), or on-site via a public lottery.

One site (Ghana) had a more complex design with two additional treatment groups (savings only, and productive asset grant only) to "unpack" those aspects of the intervention. In this paper, we are using only the group that received the pooled intervention. This is because none of the other studies systematically tried to unpack the effects, and therefore even with the full Ghana results we would have just one "data-point" and would not be able to answer the unpacking questions with anything approaching the degree of confidence that we have about the overall program effect.

The sample size used in the analysis varies from 925 households (Ethiopia) to 2,606 households (Ghana) from site to site. The overall sample size pooling all sites is 10,495 households.

Table 2 provides details by site of key experimental design features, including sample sizes; Fig. 1 provides a timeline for the typical implementation of both the program and the data collection; and figs. S1a to S1f provide a timeline for each site.

Table 2. Research design.

	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
Location	Tigray Region	Northern Region	Lempira Department	West Bengal State	Sindh Region	Cusco Department
Sample Size	925	2,606	2,403	978	1,299	2,284
Number of households chosen in randomization	458	666	800	512	660	785
Number of households accepting entry into program	458	666	800	266	660	785
Short consumption surveys	No	Yes, three rounds	Yes, five rounds	No	No	Yes, eight rounds
Spillover research design	No	Yes, GUP Villages	Yes	No	No	Yes
Multiple treatment arms	No	GUP savings, GUP no savings, SOUP matched, SOUP unmatched, Asset only	No	No	No	No
Unit of randomization	Individual household randomization	Village-level randomization, followed by individual household randomization within treatment villages	Village-level randomization, followed by individual household randomization within treatment villages	Individual household randomization	Individual household randomization	Village-level randomization, followed by individual household randomization within treatment villages
Method of randomization	One public lottery in each of 10 tabias (administrative subunits). In each tabia, 50 were selected for treatment status, 50 for control	Randomization done remotely by research team	Randomization done remotely by research team	Randomization done remotely by research team	Public lottery	Randomization done remotely by research team
Stratification variables used in block randomization procedure	10 tabia	155 villages (among those not selected as pure control villages)	40 villages (among those not selected as pure control villages)	119 hamlets	4 partner NGOs; 66 villages	43 villages (among those not selected as pure control villages)
Stratification variables to verify orthogonality, used in rerandomization procedure	None	Household size, asset ownership index, household owns business, total surface area of land owned, livestock ownership index, distance to closest market, number of compounds in village	Household size, number of children, female headed household, adult in household of prime working age (18 to 60), number of durable goods owned, income per capita, cell phone ownership	None	None	Household size, total PPI score, education level of household head, population in village, distance from nearest town
Number of months between midpoint of baseline survey and midpoint of asset transfers	6	10	15	0	11	7
Number of months between midpoint of asset transfer and midpoint of endline 1	21	18	25	23	25	29

Continued on next page

	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
Number of months between end of household visits (i.e., end of program interaction with households) and midpoint of endline 1	1	1	3	5	Varied by partner	2
Number of months between midpoint of endline 1 and midpoint of endline 2	13	12	13	15	7	12

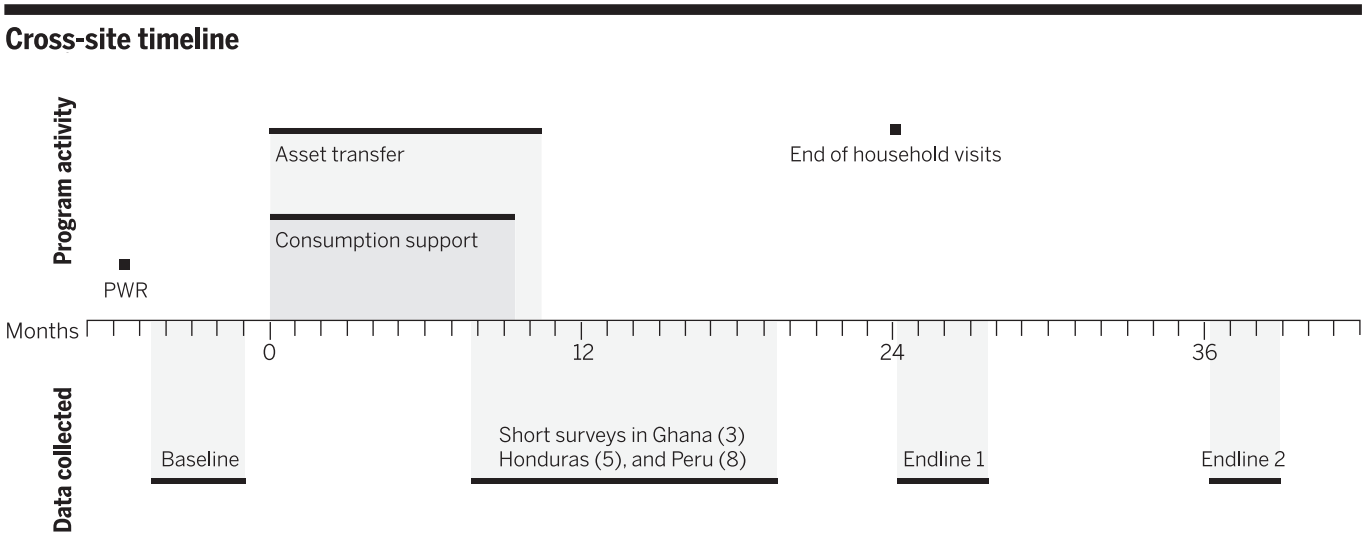


Fig. 1. Cross-site timeline. PWR, participatory wealth ranking. Only short surveys that occurred within 12 months of endline 1 are used in endline 1 analysis.

Integrity of the experimental design

Balance

Table S1b presents baseline data for the same variables and indices used as the primary outcome measures. Panel A presents the mean comparisons and t tests for equality of means. At baseline, we fail to reject at the 5% level the equality of means of treatment and control groups for any of the 10 primary outcome measures. Panel B presents similar analysis, but with a regression framework that includes fixed effects by country, and finds similar balances. The aggregate test, reported in panel C, finds that we are not able to reject equality of means across all 10 measures (p-value = 0.689). Tables S1c to S1e, present similar results for each country. Overall, the sample balance was good in every individual country.

Survey attrition

Table S1f presents an analysis of survey attrition for both endlines 1 and 2. The follow-up rate was excellent. We resurveyed 94% of baseline respondents in endline 1, and 91% in endline 2 (panel A). Panel B presents analysis on the type of people that were more likely to be resurveyed. Panel C presents a test of whether the treatment affected the type of person who completed the endline surveys, i.e., whether the treatment caused a sample composition bias. The p-values on a full set of baseline characteristics interacted with treatment are 0.75 (endline 1) and 0.17 (endline 2), thus supporting the contention that the survey attrition did not lead to a different sample frame across treatment and control groups (*17*). Tables S1g and S1h present similar results for each country. At 17%, attrition was the worst in India in endline 1; Pakistan was the worst at endline 2, at 21%. In neither country was attrition differential in the treatment group.

Compliance with treatment assignment

In all sites but one, the experimental design was strictly adhered to: No control received the program, and all treatment households received the program. The India site was the only site in which some individuals refused participation: 52% of those selected in the randomization participated in the program. According to Bandhan, the implementing organization, 35% of households declined the offer, for two unrelated reasons: First, in some villages, a section of villagers held the (erroneous) belief that Bandhan was a Christian organization trying to convert beneficiaries, and acceptance of the livestock constituted agreeing in some way to participating in Christian rituals. Second, some wives were worried that their husband would mishandle the asset and they would lose face in front of their village. A further 13% were deemed ineligible by Bandhan because they were participating in microcredit or self-help group activities. The analysis below is an "intent-to-treat" (ITT): We compare households assigned to control to those assigned to treatment, irrespective of whether they received treatment or not.

Analysis methods for pooled results

Following standard practice in the analysis of multi-site trials, we estimate a single model, with strata and country dummies. Each column of each table represents the results of a separate ordinary least squares (OLS) regression of the form

$$Y_i^k = \alpha + \beta_1 assignment_i + \beta Z_i^k + T_{\text{country}} + U_{\text{shortsurveys}} + V_{\text{stratification}} + \varepsilon_i \quad (1)$$

where Y_i^k is the outcome k of interest for either household or adult i (details of the variable constructions are presented in the supplementary

text 1 to 3), $assignment_i$ is an indicator for having been randomly selected into the program, Z_i^k is the household or adult's baseline value of the outcome variable k (coded as zero, with an indicator for missing baseline, whenever it was not available), $T_{country}$ is a vector of dummy variables for each of the countries in the study, $U_{shortsurveys}$ is a vector of dummy variables indicating whether or not the household was surveyed in a short survey round (in some countries, data were collected through both long and short surveys), and $V_{stratification}$ is the vector of all variables included in stratification in each of the six countries (*18*).

In the main analysis of the pooled sample, no adjustments are made to reflect the differences in sample sizes between countries; every observation is weighted equally. Again, this follows standard practice in the analysis of multisite RCTs. Regressions that instead weigh each country equally generate similar results. For each variable that we report, we also present the result of a test for equality of the effects across sites (which we discuss in the next subsection).

Because of the comprehensive nature of the program, a large number of outcome variables are reported. Therefore, we expect some of the variables to show significant results due to chance. To avoid overemphasis on any single significant result, we take several steps. First, following Kling *et al.* (*19*), for each "family" of outcomes, we report an index of all of the outcomes taken together, which we report in Table 3. This is our main results table. We construct indices first by defining each outcome Y_{ijl}^k (outcome k, for observation i in family j, within country l) so that higher values correspond to better outcomes. Then we standardize each outcome into a z-score, by subtracting the country control group mean at the corresponding survey round and dividing by the country l's control group standard deviation (SD) at the corresponding survey round. We then average all the z-scores, and again standardize to the control group within each country and round (*20*).

Second, given that multiple families of outcomes are being reported, we correct for the potential issue of simultaneous inference using multiple inference testing. We calculate q-values using the Benjamini-Hochberg step-up method (*21*) to control for the false discovery rate (FDR). We follow the procedure outlined in Anderson (*22*), and test α at all significance levels (1.000, 0.999, 0.998… 0.000). Our q-value is the smallest α at which the null hypothesis is rejected. It is reported in Table 3 (*23*).

Results

Pooled sample

Table 3 (both endlines), fig. S2 (endline 1), and Fig. 2 (endline 2) present an overview of the results pooled across all sites. Table 3 shows the results aggregated by "families," including q-values corrected for the fact that we are presenting the results from 10 indices (*24*).

At endline 1 (year 2 of the study, just after the end of the program in most sites), all the families of outcomes have improved in the treatment group (compared to the control group). We use two outcome measures for consumption: Per capita consumption increases by 0.12 SDs (q-value 0.001), which is equivalent to PPP US$4.55 per capita per month, or roughly 5% of control group mean of PPP US$78.80; and an index of food security increases by 0.11 SDs (q-value 0.001). An index of productive and household assets increases by 0.26 SDs (q-value 0.001). Household income and revenues increase by 0.38 SDs (q-value 0.001). There are also improvements in personal lives: Physical health improves by 0.034 SDs (q-value 0.078), and mental health improves by 0.10 SDs (q-value 0.001). Political involvement increases by 0.064 SDs (q-value 0.001), and women's empowerment by 0.046 SDs (q-value 0.049).

By endline 2 (year 3 of the study, typically 1 year after the program ended), all the effects on economic variables are still significant, and usually similar to or larger than after endline 1. It is striking that there is no evidence of mean reversion in the per capita consumption, food security, or assets. The gains in financial inclusion, total time spent working, income and revenue, and mental health have declined but are still positive and statistically significant. The gains in physical health and women's empowerment have declined and are no longer statistically significant.

Figure S2 and Fig. 2, which present the variable-by-variable results at a glance, tell a similar story: The indices are not driven by specific variables. Most individual variables show significant impacts at endline 1. At endline 2, most variables stay significant, and the various variables in the women's decision-making families and the mental health families have either declined or become not significant.

Tables S2a to S2h, contain the detailed variable-by-variable results for the entire sample.

In table S2a, we see that food consumption increases more than nonfood consumption, both in absolute value and in proportion (specifically, food consumption increases 7.5% from a control group mean of $51.60, and nonfood consumption increases 2.4% from a control group mean of $25.30). The elasticity of food consumption to overall expenditure appears to be greater than 1, a striking result given prior estimates of well below 1 (*25*). Durable goods expenditures do not increase significantly in either time period, but we do see that treatment households have more household assets than the control households in both periods (table S2c), so the expenditure variable may fail to pick up some durable goods expenditures. The consequence of the increase in food expenditure is a greater sense of food security (table S2b), which is as strong in endline 2 as in endline 1 (for example, 14% reported at least one person not eating at all for an entire day, compared to 17% in the control group; table S2b, column 3).

In table S2c, we see that households have statistically significantly more assets both in endline 1 and in endline 2. The asset index we construct in all countries is 0.26 SDs larger in endline 1 and 0.25 SDs larger in endline 2. Likewise, the effect size for productive assets (those used in household self-employment activities) does not change between endlines 1 and 2, with an effect size of 0.27 SDs at endline 1 and 0.25 SDs at endline 2. There is an increase both in household and productive assets, but the increase in productive assets is larger in both years (productive asset value increases by 15.1 and 13.6% compared to control group means of PPP US$1964 and PPP US$1576 in endline 1 and 2, respectively). Row 12 of Table 4 compares the value of the assets held by households by year 3 to the value of the asset that was transferred to them. The impact of the program on asset values is lower than the cost of the assets. However, the program impact on asset holdings is stable from year 2 to year 3 (Table 3), so after the households made an initial adjustment to asset holdings, there was no further decline.

The increase in asset holding does not come at the expense of more borrowing or less savings. Instead, we see in table S2d large increase in savings in both endlines (PPP US$151, or 155.5% of control mean in endline 1, and PPP US$75, or 95.7% of control means in endline 2). Savings was mandatory during the first year in many sites, so it is not entirely surprising that we see an increase at endline 1. But continued savings was not required after the program, and the increase in net savings is still large.

These productive assets are being put to use: Adult labor supply increases by 17.5 min per adult per day (10.4% increase over control households) at endline 1, and 11.2 min (6.1% increase) at endline 2 (table S2e). The increase is concentrated on livestock and agricultural activities, consistent with the assets chosen by most people. More assets and more labor translate into increased revenue from livestock (table S2f, column 1) (*26*) and net income from agriculture (column 2). At endline 1, the revenue from livestock is 41.6% larger, compared to a control group mean of PPP US$73.50). At endline 2 it is 37.5% larger, compared to a control group mean of PPP US$80.60. The households also feel better off economically: 0.33 points improvement on a scale of 1 to 10 at endline 1 (control group mean = 3.74), and 0.30 points improvement at endline 2 (control group mean = 3.65). All of the gains to income and revenue persist 1 year after the end of the program, including the increase in self-reported economic status.

Table S2g presents the detailed health and mental health results. The only significant positive impact on physical health seen at either endline at the 5% level is on the activities of daily living score at endline 1. At endline 1, the mental health index is 0.10 SDs higher, driven by the overall self-reported happiness and lack of symptoms of mental distress (*27*). By endline 2, the positive impact on the mental health index has declined to 0.071 SDs, but it remains significantly positive and continues to be driven by both self-reported happiness and lack of stress. This minor decrease in the treatment effect may be another instance of the well-known "hedonic treadmill" (*28*).

Table S2h presents results on political and social empowerment, and women's empowerment

Table 3. Indexed family outcome variables and aggregates.

	Endline 1			Endline 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Indexed outcomes	Standardized mean treatment effect	*q*-value for all 10 hypotheses	*F*-test of equality of coefficients across sites, with *q*-values	Standardized mean treatment effect	*q*-value for all 10 hypotheses	*F*-test of equality of coefficients across sites, with *q*-values
Total per capita consumption, standardized	0.122*** (0.023)	0.001	3.207 0.009	0.120*** (0.024)	0.001	5.307 0.001
Food security index (five components)	0.107*** (0.022)	0.001	1.670 0.139	0.113*** (0.022)	0.001	2.405 0.050
Asset index	0.258*** (0.023)	0.001	14.26 0.001	0.249*** (0.024)	0.001	23.90 0.001
Financial inclusion index (four components)	0.367*** (0.030)	0.001	55.33 0.001	0.212*** (0.031)	0.001	10.70 0.001
Total time spent working, standardized	0.090*** (0.018)	0.001	7.520 0.001	0.054*** (0.018)	0.004	2.644 0.038
Incomes and revenues index (five components)	0.383*** (0.036)	0.001	12.05 0.001	0.273*** (0.029)	0.001	5.82 0.001
Physical health index (three components)	0.034* (0.019)	0.078	3.825 0.003	0.029 (0.020)	0.159	0.776 0.630
Mental health index (three components)	0.099*** (0.022)	0.001	5.189 0.001	0.071*** (0.020)	0.001	1.781 0.142
Political Involvement index (four components)	0.064*** (0.018)	0.001	4.176 0.002	0.064*** (0.019)	0.002	2.624 0.038
Women's empowerment index (five components)	0.046** (0.023)	0.049	1.803 0.121	0.022 (0.025)	0.385	0.469 0.800

Notes: 1. Results presented are mean standardized intent-to-treat estimates, including country dummies and controls for every variable used in both block stratification and in rerandomization procedures. All indices are standardized with respect to the control group in that same time period. 2. Dummy variables are included in endline 1 regressions for wave of data, i.e., for whether observation was from endline 1 or one of the short surveys conducted prior to endline 1. 3. See supplementary text 2 for the components of each index 4. Endline 1 was conducted immediately following the end of the household visits, which was typically 2 years after the transfer of productive assets. 5. Endline 2 was conducted 12 months after endline 1, i.e., 1 year after the end of all program activities. 6. Indices measured at the household level are total per capita consumption, food security, assets, financial inclusion, and incomes and revenues. For these indices, our sample size ranges from 9,613 to 9,785 for endline 1, and from 9,482 to 9,508 at endline 2. Use of time, physical health, mental health, and political involvement are asked of adults in each household (normally one or two adults, but as many as seven in India). Sample sizes for these indices range from 12,493 to 15,662 at endline 1 and from 14,051 to 15,136 at endline 2. Pakistan did not include a mental health module in endline 1. India did not include a women's empowerment module in endline 2. 7. For both household- and adult-level indices, standard errors are clustered at the unit of randomization.

within the household. Beneficiaries, who are at the outset often marginalized within their village, become more likely to be involved in political activity (except voting) and village-level actions. This improvement is true both immediately after the program ends and 1 year later. At endline 1, treatment women report having a greater say in decisions within the household related to health expenditures and home improvements. However, this gain in empowerment does not persist over time.

In table S3, we present bounds for our treatment effects, depending on different assumption with respect to attrition, using Horowitz-Manski-Lee bounds (*29*, *30*). The conclusions are robust to this exercise, with all lower bounds except that for women's empowerment significantly positive at endline 1.

Country-by-country variation

There are too many countries and too many variables to comment on the country-by-country and variable-by-variable results in detail, though the tables are all available in the supplementary materials. Figure S3 (endline 1) and Fig. 3 (endline 2) have a format similar to that of fig. S2 and Fig. 2, but they present the country-by country results for the summary indices. Tables S4a through S4f present the impacts on the 10 indexed family outcomes, one table per country. Tables S5a-1 through S5h-2 present the impacts on each of the components in each of the countries, one table per family of outcomes per endline. Here, we highlight some particularly relevant information from this analysis.

The first and most important point is that the results are not driven by any one country. The differences across countries can be seen in fig. S3 and Fig. 3. We present tests for the hypothesis that the results are the same for all countries for each outcome variable. The hypothesis is rejected for almost all pooled outcomes (Table 3), which suggests that there is significant site-by-site variation (and enough data to pick it up), which would be important to study in future work. However, in endline 1, the program appears to have positive impacts on most indices for most countries (tables S4a to S4f). An exception is

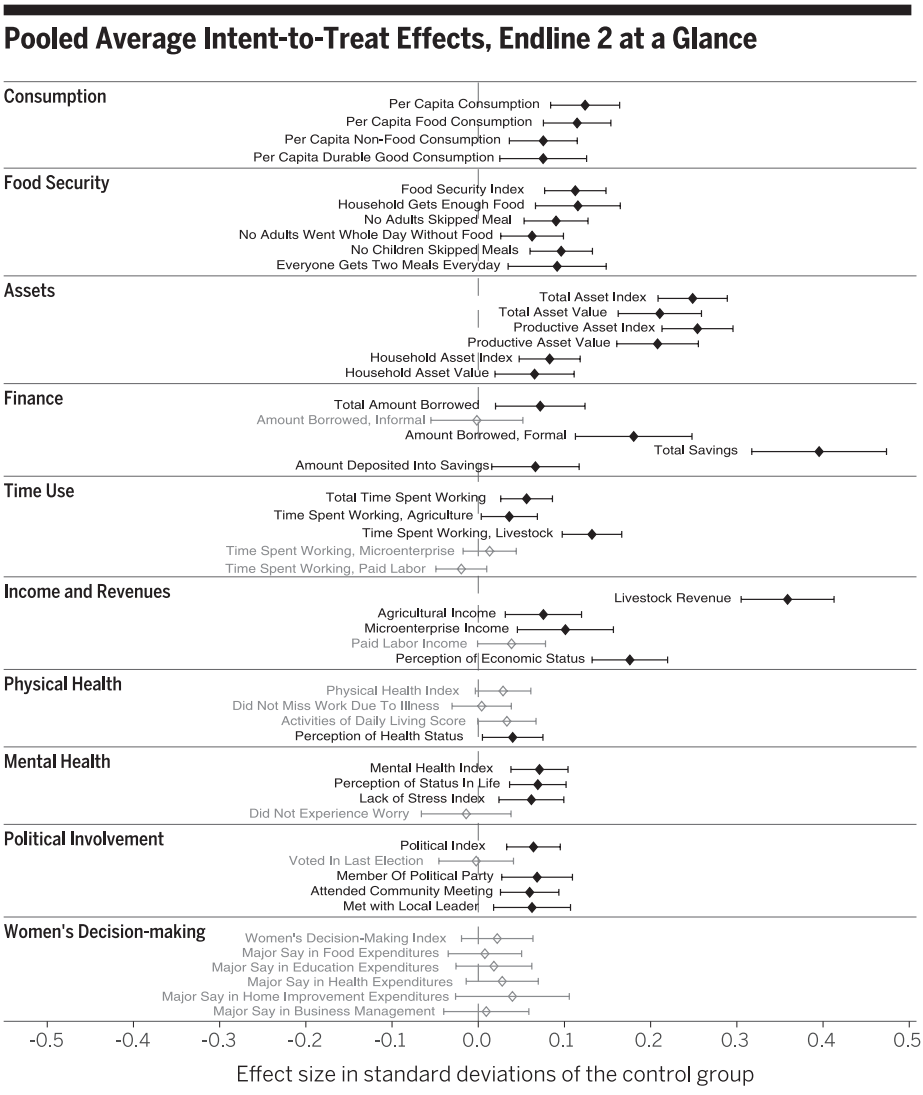


Fig. 2. Pooled average intent-to-treat effects, endline 2 at a glance. This figure summarizes treatment effects presented in tables S2a to S2h. Treatment effects on continuous variables are presented in SD units. Each entry shows the OLS estimate and 95% confidence interval for that outcome. (◆) Statistically significant, 5% level; (◇) not statistically significant, 5% level.

Peru, where we see 3 results out of 10 statistically significant at the 5% level. In endline 2, four of the countries (Ethiopia, Ghana, India, and Pakistan) continue to have statistically significant and positive impacts on most variables, but Honduras and Peru have weaker results, with positive and statistically significant impacts on 3 out of 10 and 4 out of 10 families of outcomes before multiple hypothesis adjustments, respectively (and Honduras also has a negative, and statistically significant prior to multiple hypothesis adjustment, impact on assets).

Turning to the key variables, the gains in per capita consumption, for example, are statistically significant for both endlines in every country except Honduras and Peru. However, we do find a statistically significant increase in food consumption for Honduras in endline 1 and Peru in endline 2 (tables S5a-1 and S5a-2). Likewise, there is an increase in livestock revenues (livestock was the most frequently chosen asset in all sites) in all countries by endline 2. There is significant improvement in assets in all countries except Honduras (where it actually declines by endline 2). When looking at the variables individually, some results are different from country to country, no doubt partly due to local specificities and probably partly due to pure luck, but the overall bottom line is that the program appears to be effective in most places. Even in Peru, where we see gains on fewer variables than in other countries, the gains in food expenditures per capita, assets, livestock revenues, physical health, and mental health are all positive and significant.

Second, although it is dangerous to rationalize the Honduras results ex-post, there is a relatively simple explanation for the pattern of results we observe, with generally positive results in endline 1 declining by endline 2. Most households were given chickens. In both endlines we do see an increase in revenue coming from chickens, as well as a significant increase in food consumption. However, a large fraction of the chickens died due to illness. By the time households were interviewed at endline 2, the households had lost most of their productive asset (leading to a negative and statistically significant impact on the asset index by endline 2) and were not consuming more.

Third, the India results, which come from West Bengal, an area of India that is directly abutting Bangladesh and shares a language and a culture, are strikingly similar, down to most details, to the results in the RCT of the impact of BRAC reported in Bandiera *et al.* (*6*). In particular, as they do, we find that there is an increase in nonagricultural, nonlivestock income by endline 2 in West Bengal (table S5f-2). None of these were promoted through the program, yet they materialized as the household's well-being improved (*31*). This result suggests a pathway—income diversification—through which the results might persist over time.

Fourth, the strongest positive results across the board are obtained in Ethiopia. The Ethiopia case is interesting, because it is the one country where all of the control group also received the basic consumption support that, in other sites, is only provided to the treatment groups (in Peru, half of the control group was also on a government cash transfer program, Juntos). Because it is only one country, we have no counterfactual to what would have happened in Ethiopia if the control group had not received consumption support, but this design at least tells us that the consumption support on its own is not responsible for the entire impact of the program. Note, however, that the productive asset transfer in Ethiopia (equivalent to 7.98 goats) was also larger than in Ghana (6.00 goat equivalents), India (6.53 goat equivalents), or Pakistan (3.75 goat equivalents), so to the extent that assets are liquid, the larger asset transfer in Ethiopia may have compensated for the difference in consumption support.

Effects on distribution of outcomes

Table 5 shows quantile regression estimates at the 10th, 25th, 50th, 75th and 90th percentiles of the distribution of the outcomes. There are several notable results. First, we see positive and significant impacts on income, consumption, and assets, at all tested quantiles. This is encouraging, in that it shows that the program did not push the poorest toward an activity that they did not have the means to manage successfully. Second, for the other variables, the pattern of results is what standard theory would predict. For example, we see impacts on food security only toward the bottom (at the 25th percentile): Those are the households that frequently miss meals and thus likely use any income gains to buy more food. On the other hand, we see impacts on financial inclusion only for the top quantiles (median and above at both endlines): If either access to credit or savings requires meeting some threshold of resources, the poorest of the

Table 4. Cost-benefit analysis.

	Panel A: Program costs per household, USD PPP 2014	Ethiopia	Ghana	Honduras	India	Pakistan	Peru
	Direct transfer costs	1228	680	724	700	2048	1095
(1)	Asset cost	1228	451	537	437	1043	854
	Food stipend	0	229	187	263	911	241
	Total supervision costs	1900	2832	1633	407	–	3357
	Salaries of implementing organization staff	347	1994	801	297	–	2477
	Materials	33	119	112	1	–	55
	Training	850	44	121	19	–	111
	Travel costs	174	293	210	17	–	55
	Other supervision expenses	496	382	388	73	–	660
	Total direct costs	3127	3513	2356	1107	4680	4452
	Start-up expenses	43	133	104	38	–	45
	Indirect costs	421	1026	209	112	470	462
	Total costs, calculated as if all incurred immediately at beginning of year 0	3591	4672	2670	1257	5150	4960
(2)	Total costs, inflated to year 3 at 5% annual discount rate	4157	5408	3090	1455	5962	5742
	Exchange rate to PPP adjustment scalar	3.41	2.19	1.90	3.52	4.44	1.84
	Panel B: Benefits per household, USD PPP, all values inflated or deflated to year 3 at 5% annual social discount rate						
(3)	Year 1 annual nondurable consumption ITT, assuming treatment effect equal to year 2	451	293	66	344	613	339
(4)	Year 2 annual nondurable consumption ITT treatment effect	451	293	66	344	613	339
(5)	Year 3 household asset ITT treatment effect	63	15	−20	6	7	37
(6)	Year 3 nondurable annual consumption ITT treatment effect	424	332	−218	251	451	263
(7)	Year 4 onward total consumption ITT treatment effect, assuming year 3 gains persist in perpetuity	9417	6241	−6011	5354	8994	7402
(8)	Total benefits: (3) + (4) + (5) + (6) + (7) = (8)	10805	7175	−6118	6298	10678	8380
(9)	Year 3 productive asset ITT treatment effect	851	118	32	171	163	59
(10)	Year 3 savings balance ITT treatment effect	272	11	32	9	7	45
	Panel C: Benefit/cost ratios						
(11)	Total benefits/total costs ratio: (8)/(2) = (11)	260%	133%	−198%	433%	179%	146%
(12)	Increase in asset value in year 3 (Household, productive and financial)/cost of asset transfers: [(5) + (9) + (10)]/(1) = (12)	97%	32%	8%	43%	17%	16%
(13)	Increase in asset value/transfers, 10th percentile	56%	5%	−3%	1%	2%	7%
(14)	Increase in asset value/transfers, 25th percentile	72%	12%	8%	10%	7%	8%
(15)	Increase in asset value/transfers, 50th percentile	85%	20%	15%	23%	15%	7%
(16)	Increase in asset value/transfers, 75th percentile	123%	29%	20%	58%	45%	16%
(17)	Increase in asset value/transfers, 90th percentile	175%	37%	32%	131%	52%	7%
	Sensitivity analysis						
(18)	Internal rate of return (IRR)	13.3%	6.9%	–	23.4%	9.5%	7.5%
(19)	Annual rate of dissipation of the treatment effect such that costs = benefits	10.3%	1.8%	–	31.1%	5.0%	2.6%
(20)	Benefit/cost ratio, at discount rate of 7%	182%	93%	−132%	306%	127%	102%
(21)	Benefit/cost ratio, at discount rate of 10%	124%	63%	−84%	211%	88%	69%

Notes: 1. Costs: The implementing partner in Honduras initially allocated all start-up costs and indirect costs into their direct supervision line items. We assume that 5% of costs were committed to start-up and 10% to indirect costs, while preserving the total costs equal to the organization's full budget for the project. In Pakistan, there were five implementing partners, each with different allocations of the nondirect costs. The total represents the average across these organizations. Staff costs associated with the selection of the beneficiaries (identifying the district and poorest individuals, and screening are included in the staff costs. The (nonstaff) costs associated with the identification process are included in "other supervision expenses." In India, note that the compliance rate was 52.0%; the cost-benefit conservatively uses the per-person cost of those who received the program (rather than the total costs divided by the number of all people selected in the randomization). 2. Benefits: In India and Ghana, individuals do not provide an estimate of the value of all assets. We use the relative value of assets across sites and the average purchase prices available in each country (e.g. goats and cattle in both sites) to provide an estimate of the asset ITT. In India and Pakistan, we do not directly ask about the savings balance. We use an OLS regression from the other four countries, with savings balance as our dependent variable and cumulative deposit amount as our independent variable, to predict the ITT of savings balance in those two countries. We calculate nondurable consumption equal to the total of columns (2) and (3) in tables S5a-1 and S5a2 (i.e., total consumption less durable good expenditures) multiplied by the average household size in the country times 12. The average household sizes used in Ethiopia, Ghana, Honduras, India, Pakistan, and Peru are 5.59, 8.34, 5.88, 3.96, 6.29, and 5.20 in endline 1, respectively, and 5.59, 8.48, 5.91, 3.79, 6.58, and 5.24 in endline 2, respectively. 3. Benefit/cost: To estimate the break-even dissipation rate (i.e., the rate of decline of the impact on consumption from one time period to the next), we calculate the net present value of consumption in perpetuity beginning in year 4 with the equation {[*ITT consumption* × (1 − *dissipation rate*)]/1.05}/(*discount rate* + *dissipation rate*). We then solve for the level of dissipation such that the net present value of the costs equals the net present value of the benefits. We do not offer a calculation for Honduras because the costs exceed the benefits even when assuming there is no dissipation. 4. See supplementary text 5 for more details on the cost-benefit calculations. 5. Table S7 replicates Panel A of this table but using exchange rate conversions rather than PPP conversions

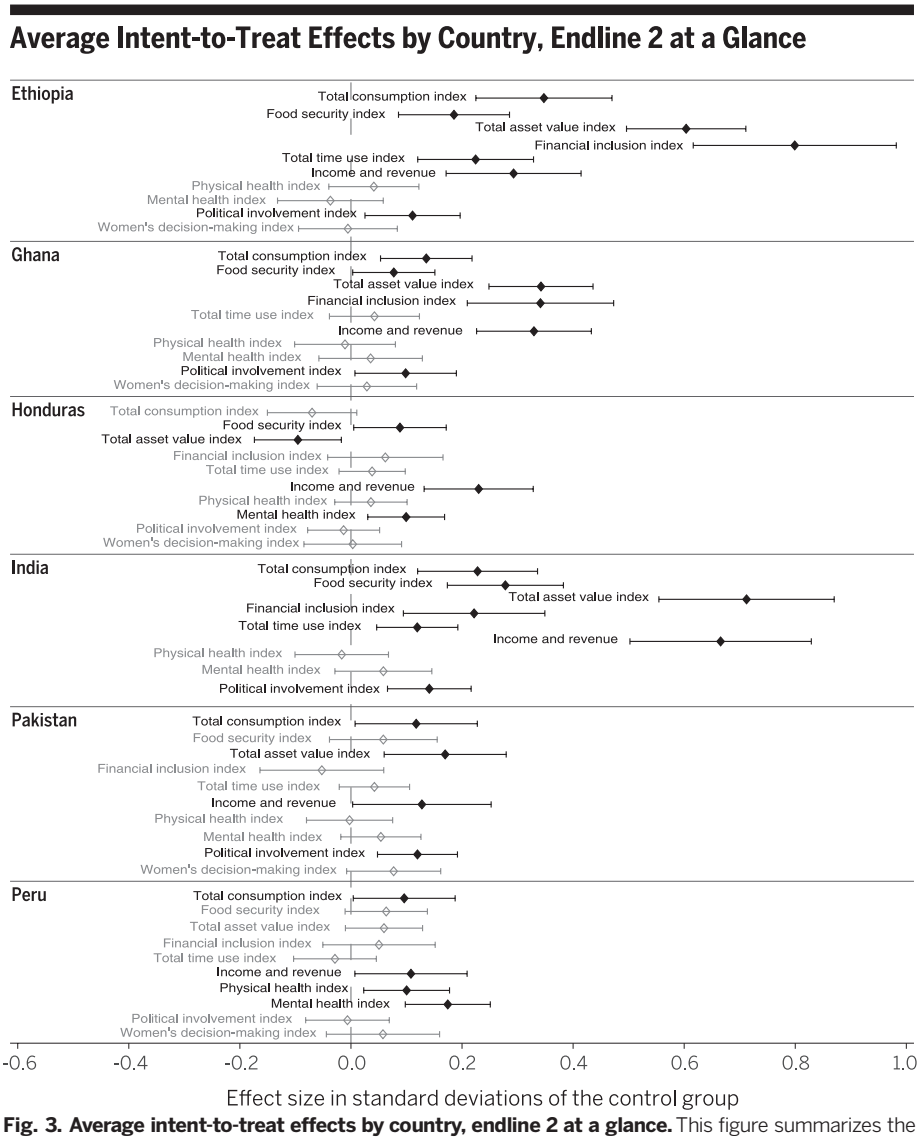


Fig. 3. Average intent-to-treat effects by country, endline 2 at a glance. This figure summarizes the treatment effects presented in tables S3a to S3f. Treatment effects are presented as *z*-score indices, standardized to the control group at endline 2. Each entry shows the standardized index outcome and its 95% confidence interval. (♦) Statistically significant, 5% level; (◇) not statistically significant, 5% level.

poor may not have met that threshold even with the program. Third, the effects on consumption per capita and the income and revenues index are all increasing with the quantiles: for example, at endline 1, the 10th percentile of consumption (income and revenue index) increases by 0.027 SD (0.005 SD), whereas the 90th percentile increases by 0.491 SD (0.079 SD). Finally, we do see much larger asset growth at higher quantiles (0.038 SD for the 10th quantile versus 0.357 for the 90th quantile).

Are spillovers biasing the results?

In supplementary text 4 and tables S6a and S6b, we examine spillover results in Ghana, Honduras, and Peru. These three sites employed a randomization at both the village and household levels to permit comparisons of control individuals in treatment villages to control individuals in control villages. Overall, these results suggest that neither externalities nor general equilibrium effects within villages substantially affect our outcomes. This finding implies that it is appropriate to pool the control households in treatment villages with the households in control villages to form the control group.

Discussion

The experiment, conducted in six countries on three continents, shows that the ultrapoor Graduation program improves the lives of the very poor along many dimensions. The program's primary goal, to increase consumption, is achieved by the conclusion of the program and maintained 1 year later. Furthermore, the pattern of impacts on intermediate and downstream outcomes accords with the theory of change: Productive assets, income, and revenue go up. Although results vary across countries, the general pattern of positive effects that persist for at least a year after the program concludes is common across all countries, with weaker impacts in Honduras and Peru.

Cost-benefit analysis

Naturally the benefits should not be considered without also considering the costs. Table 4, panel A presents costing details, broken down by direct costs (direct transfer and supervision costs), start-up expenses, and indirect costs (including local and international overhead costs). The total program costs for the full duration of the program (inflated to year 3 equivalent PPP dollars, using 5% as the social discount rate, range from PPP US$1455 per household (India) to PPP US $5962 (Pakistan). We use 5% as the social discount rate to harmonize with the joint World Bank and International Monetary Fund policy (*32*), but also calculate internal rates of return and show sensitivity to 7 and 10%. There is no single driver of costs to explain the differences; some of it can be attributed to in-country operating cost differences and some is presumably due to variations in the actual program design. Peru, for example, is a much richer country than Pakistan, so the wages paid to the implementing staff were a lot higher. It is not possible to precisely assign labor costs to specific activities; however, the majority of supervision costs in each country are likely attributable to the household visits and training activities. The asset costs and food stipends, by contrast, required little labor to distribute.

Table 4, Panel B summarizes the consumption gains and asset value changes attributable to the program, all inflated to year 3 equivalent PPP dollars. We assume that the (unmeasured) year 1 ITT effect on per capita consumption is equal to that estimated for year 2, and we assume that the estimated impact on year 3 consumption continues indefinitely into the future (we then relax this assumption, below, as a sensitivity check). The overall impact of the program on consumption expenditure, reported in row 8, is the sum of the impact on the year 3 stock of household durables and the total impact on each year's nondurable consumption (in year 3 equivalent dollars). Rows 9 and 10 of Panel B also report the impact of the programs on the stock of productive assets and savings.

As noted previously, the increase in assets held by the households is lower than the value of the asset in all countries but Ethiopia. On average, households have drawn down part of the asset transfer in the first year, but there is no further decline between year 1 and year 2, and the consumption gains (the final objective of the programs) persist over time. The decline in asset holding in the first year, followed by a stable pattern in both assets and consumption, is somewhat surprising, as economic theory would suggest a slower adjustment to a steady-state level of assets (even if the initial transfer was larger than the optimal steady-state level of assets). We may capture imperfectly some informal assets or liabilities (such as debt or loans to or from other households in the village, which may be labeled as gifts or alms). We also do not capture the value of human capital, which has increased as a result of better nutrition,

Table 5. Quantile treatment effects, indexed family outcomes.

Indexed outcomes	Endline 1					Endline 2				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Total per capita consumption, standardized	0.065*** (0.014)	0.067*** (0.013)	0.045** (0.019)	0.099*** (0.032)	0.140** (0.058)	0.053*** (0.013)	0.061*** (0.014)	0.079*** (0.019)	0.105*** (0.030)	0.206*** (0.063)
Food security index (five components)	0.013 (0.029)	0.073** (0.034)	0.000 (0.000)	0.000 (0.006)	0.000 (0.002)	0.000 (0.020)	0.032** (0.015)	0.000 (0.010)	0.000 (0.006)	0.000 (0.002)
Asset index (productive and household assets)	0.038*** (0.007)	0.072*** (0.010)	0.172*** (0.017)	0.288*** (0.028)	0.357*** (0.064)	0.023*** (0.007)	0.062*** (0.009)	0.135*** (0.014)	0.251*** (0.026)	0.329*** (0.054)
Financial inclusion index (four components)	0.000 (0.001)	0.000 (0.001)	0.153*** (0.010)	0.322*** (0.023)	0.587*** (0.072)	0.000 (0.002)	0.000 (0.002)	0.033*** (0.011)	0.213*** (0.028)	0.418*** (0.095)
Total time spent working, standardized	0.000 (0.006)	0.000 (0.005)	0.084*** (0.018)	0.078** (0.034)	0.000 (0.028)	0.000 (0.006)	0.000 (0.005)	0.038* (0.021)	0.045* (0.027)	0.022 (0.039)
Incomes and revenues index (five components)	0.027*** (0.005)	0.053*** (0.008)	0.098*** (0.014)	0.314*** (0.034)	0.491*** (0.079)	0.001 (0.006)	0.029*** (0.009)	0.171*** (0.020)	0.307*** (0.034)	0.538*** (0.070)
Physical health index (three components)	0.000 (0.033)	0.000 (0.024)	0.012 (0.019)	0.000 (0.007)	0.000 (0.006)	0.000 (0.035)	0.000 (0.022)	0.000 (0.016)	0.000 (0.008)	0.000 (0.006)
Mental health index (three components)	0.108*** (0.038)	0.005 (0.025)	0.112*** (0.027)	0.082*** (0.026)	0.000 (0.015)	0.000 (0.020)	0.000 (0.018)	0.000 (0.018)	0.000 (0.011)	0.000 (0.009)
Political involvement index (four components)	0.000 (0.014)	0.000 (0.010)	0.000 (0.007)	0.000 (0.008)	0.000 (0.006)	0.000 (0.004)	0.000 (0.003)	0.000 (0.010)	0.000 (0.009)	0.000 (0.006)
Women's empowerment index (five components)	0.000 (0.014)	0.000 (0.009)	0.000 (0.008)	0.000 0.000	0.000 (0.023)	0.000 (0.014)	0.000 (0.015)	0.000 (0.008)	0.000 (0.008)	0.000 (0.009)

Notes: 1. Regressions include baseline controls, country dummies, and controls of the variables used to ensure balance in rerandomizations at baseline. They do not include block stratification variables. 2. The 0's as coefficients reflect the fact that several of our indices relied either partially our entirely upon categorical variables. Although the distributions of results differ at some points, at each percentile measured there is no treatment effect 3. We perform a Komolgorov-Smirnov test for the equality of distributions between treatment and control households for each indexed outcome in both endlines. The *p*-value for each of these 20 tests is 0.000.

physical and mental health: Spending on better food and needed health expenditures early in the program may have been a valuable investment.

The ultimate goal of the program is to durably increase consumption, not merely to increase asset holding. Using total consumption as the measure for benefits, the total benefit-cost ratios presented in row 11 indicate that with the exception of Honduras, the programs all have benefits greater than their costs (ranging from 133% in Ghana to 433% in India).

We explore the sensitivity of this conclusion to some of our crucial assumptions. First, we calculate the internal rate of return, to assess at what social discount rate costs equal benefits. They are 13.3% (Ethiopia), 6.9% (Ghana), not applicable (Honduras), 23.4% (India), 9.5% (Pakistan), and 7.5% (Peru). Second, we calculate in row 19 the rate at which nondurable consumption must dissipate after year 3 (rather than persist into the future) in order for benefits to equal costs. Third, in the subsequent two rows, 20 and 21, we show the sensitivity of the benefit-cost ratio to alternative social discount rates of 7 and 10%. Benefits continue into the future while the costs are front-loaded, so the benefit-cost ratios decline with increases in the assumed social discount rate. See supplementary text 5 for details on the cost-benefit analysis calculations.

Mechanisms

As mentioned, the results are similar to the positive results of the evaluation of the BRAC program in Bangladesh (*6*). Two other studies of cash transfers and support for self-employment, both in Uganda, find similar results. Blattman *et al.* (*33*) find that a program that provided a $150 grant (PPP US$401) toward a nonfarming self-employment activity along with training and follow-up guidance to very poor women in conflict-affected regions increased consumption, cash earnings, labor supply, and nonfarm self-employment. Blattman *et al.* (*34*) find that a program that provided both training and support and a cash grant to youth increased business assets by 57%, work hours by 17%, and earnings by 38%. The programs that we studied differ from those reported on in (*33*) and (*34*) on a few dimensions: choice of sample frame (representative ultrapoor, versus unemployed young men or poor women); the level of intervention [household, versus group-level investments as in (*34*)]; and the integration of other components (health and access to savings). Nevertheless, these studies add to an emerging picture from a variety of countries that these types of programs can be effective.

Although we see impacts on all outcomes, more work is needed on the mechanisms that underlie the positive impacts. The core fact is that a time-limited big push led to a sustained increase in consumption and income. One common way to think about the effect of a big push is through the lens of the large, primarily theoretical, literature on poverty traps (*35*). In such models, the combination of constraints and incentives faced by the poor act to keep them in place, ensuring that any small improvement in their well-being quickly dissipates. Only a big push that appreciably relaxes those constraints can set off a virtuous cycle where the beneficiaries move to an entirely different trajectory.

The fact that the effects of the program seem durable supports the interpretation that the program unlocked a poverty trap. Nevertheless, the average effects are not very large and do not correspond to our intuitive sense of what it would mean to be liberated from the trap of poverty. There are several possible ways to resolve this tension:

First, it could be that there is no trap—but rather what one might call a "poverty flat," a world in which small changes persist but neither unleash continued improvement, thus leading to large longer term changes, nor dissipate rapidly.

Second, it is possible that this particular trap is small—the beneficiaries have gotten out of it, only to join the broader mass of the poor, who might be in some other, bigger, trap.

Third, it is worth recalling that the theory predicts that the effect of a push will be heterogeneous, unless the push is simply enormous. Those who are closer to the edge of the trap will exit, but the rest will just slowly fall back in. Perhaps this is what happened—the heterogeneity in the impacts that we see across the distribution lends some support to this hypothesis. Even among the very poor households targeted by these programs, the impacts on income and revenues and consumption, though positive everywhere, are lower at the bottom of the distribution. Because everyone was offered the same menu of assets, under the standard assumptions of constant or decreasing returns to the assets and homotheticity of preferences, we would expect those impacts to be either constant or decreasing. Instead, it appears that the poorest of the poor either have a lower return to the asset, or that they chose to consume more of it, or both. The differences in terms of final asset accumulation are very large: by endline 2, the point estimate of the impact of the program at the 90th percentile of the asset index is more than 10 times that at the 10th percentile.

Fourth, another source of heterogeneity, the level of patience or return on investment, could also help to explain why the average impact is both durable and yet not very large. The more patient or productive would use the asset transfer as a springboard to accumulate more assets and permanently be on a different consumption trajectory, whereas the others would sell off some part of the transferred assets to consume more than they earn, and perhaps eventually end up where they started. In rows 13 to 17 of Table 4, we use quantile treatment effects to generate the total gain in assets at different quantiles and present them relative to value of the original transfer. The ratio of the asset gain to the cost of the transfer is less than 1 at all tested quantiles in every country except Ethiopia (above 1 for the 75th and 90th percentiles) and India (above 1 for the 90th percentile), suggesting that the general pattern of eating into assets holds at every quantile. Therefore, we do not find strong evidence for this kind of heterogeneity.

But what would be the specific nature of a trap? One standard narrative for a poverty trap essentially says that poor people remain poor because they cannot afford enough food to make them strong enough to be productive (*36*). This theory has been discounted in recent years on grounds of empirical plausibility—essentially most poor people can afford to spend more on food if that were a priority for them (*37*). However, this may be a case where that theory does apply, at least to some participants in the program, because these people are poorer than most poor people and may actually not be able to afford enough food (Table 1 reports the daily per capita calories that could be purchased if baseline expenditures were allocated solely to staple grains). As noted, for the very poor, we do see large increases in food security. Moreover, the elasticity of food consumption is greater than 1 in the overall experimental population. However, this is driven by the food expenditure responses in Ghana, Honduras, and Peru. In the three other countries, the proportional increase in nonfood consumption is either similar to or greater than the proportional increase in food consumption, and we see a persistent effect there as well. Moreover, we see even larger and persistent impact even at the quantiles where there is no impact on food security (although it could still be that they are eating more nutritional food). So nutrition cannot be the whole story, although it may well be a part.

An alternative view of the poverty trap emphasizes underinvestment by the poor, either because they are unable to borrow enough to be able to make the necessary investment or because they find it too risky (*35*, *38*, *39*). For the poorest within our study, we do not find an impact on financial inclusion, and we find a weaker impact on assets. This is consistent with the need to satisfy some asset threshold before being eligible for credit, one of the key ingredients for a credit-based poverty trap. Again, however, the evidence for the existence of such traps is not very strong. There is a growing body of evidence on microcredit that was intended to improve credit access among the segment of the population only slightly less poor than our targeted group. For example, Banerjee *et al.* (*40*) review six randomized studies of microcredit in six different countries (*41*–*46*) and conclude that although microcredit loans sometimes lead to an increase in business activity, the effect on average business profits is much more muted, and there is no effect of an impact on consumption over a 1- to 3-year time period. That is, for the average poor person, better access to microcredit does not seem to generate the kind of sustained consumption gains that we see with this program, suggesting that credit alone is not the explanation.

The programs that we analyze are different from microcredit in multiple ways. Here, households did not need to repay. This might have encouraged them to take more risks and genuinely invest themselves in the activity. Or it could be the training and personal encouragement that produced this effect on their behavior. Or these people may be in a different economic position—the microcredit borrowers already have an occupation and an income and are merely trying to expand, not start a new activity. The participants in the Graduation programs are starting new activities, more or less from nothing. These are all important possibilities that deserve exploration.

But perhaps we need to go beyond these standard theories. There are now behavioral theories of poverty traps that give an important role to positive expectations of the future (*47*, *48*). We do see some improvement in the self-reported well-being of the beneficiaries, which, at endline 1, are visible at all level of the distribution except for the 90th percentile. Much more detailed psychological measurement would be necessary to fully understand this result and its underlying mechanisms. Perhaps this program worked by making the beneficiaries feel that they mattered, that the rest of society cared about them, that with this initial help they now had some control over their future well-being, and therefore, the future could be better.

These positive results leave us with a number of important questions. First, is it better to deliver physical assets and support, rather than pure cash transfers? There is evidence—from an RCT evaluation of the GiveDirectly program in Kenya, which transferred on average PPP US $720 to poor households, either monthly or in one lump sum—that pure cash transfers also have positive impacts on consumption, food security, asset holdings in the short run (including productive assets), and on psychological well-being (*49*). Similarly, de Mel *et al.* (*50*) find that a cash (or in-kind) transfer to existing self-employed individuals in Sri Lanka has a persistent positive effect on self-employment profits 4.5 to 5.5 years later. Because it is cheaper and easier to just deliver cash rather than physical assets and training, and the initial consumption increases from Kenya seem to be higher than what we observe after 2 and 3 years, it would be useful to have a direct comparison of the effects of these programs. The Ghana experimental design does include a comparison of the Graduation program to merely an asset transfer, and the results are forthcoming.

However, the Kenya results are unfortunately not quite comparable, because the time to follow-up was much shorter (4 months). The Kenya study did employ random variation in survey timing to try to examine persistence of the impact, and found that the estimated treatment effect was reduced by about half from 1 month after the transfer to 7 months; however, this reduction was not statistically significant. We observe no decline in the gain in consumption per capita almost 3 years after the asset transfer. If the effects of one-time transfers dissipate rapidly in one case and are permanent in the other, this obviously has major consequences for the comparative cost-benefit analyses of the two programs. The evolution of the impacts over time over a longer horizon thus needs to be further explored, both for pure cash transfer programs and for these broader programs.

Second, how important was the training and coaching as a component in the full intervention? This is a particularly important component

to test, because its costs are on average twice that of the direct transfer costs, and because operating at scale requires quality hiring, training, and staff supervision. As discussed above, we do not have experimental variation with which to test this question. Evidence from elsewhere suggests that the household visits, which are a large expenditure, may not be a cost-effective component. In Blattman *et al.* (*33*), for example, variation between zero and five household visits did not generate, after 9 months, large differences in income outcomes (but did lead to higher investment). Furthermore, a meta-analysis of self-employment training programs has found mixed but rarely transformative impacts from training (*51*).

This brings us to the next key question: How long will these results persist? This will not be known until some participants are followed for a longer period of time, but there are a number of encouraging signs. First, the effect on consumption does not decline over time as one would have expected had the program not led to long-term increases in income. Similarly, the increase in consumption was not generated by merely spending down the asset provided (*52*). Second, treatment households have more productive assets and have increased their labor supply 1 year after the program ends, and in some countries have diversified out of the original asset that was provided. Finally, in Bangladesh (*3*), households were followed for 2 more years after the end of the period of support, and the study continues to find robust impacts on consumption, productive assets, and earnings.

Another issue is the potential for externalities or general equilibrium effects, both positive and negative, from the program. Transferring (often) the same productive asset to many households in a small village may generate a negative externality on other asset owners, if, for example, the transfers result in a fall in the price of cows or milk. On the other hand, the benefits that accrue to the treatment households may be shared with others, as has been observed from a conditional cash transfer program in Mexico (*53*). It is worth pointing out that the program is designed to serve few people (the poorest) within each village, and in that sense, the current design probably picks up a fair share of the possible externalities. In endline 2, the evidence from the three sites where randomization allowed the examination of spillover shows no effects on primary economic outcomes such as consumption and income, and no significant effects at the 5% level on any variable after accounting for multiple hypothesis testing.

These questions will become ever more important as these programs scale. The programs studied here were implemented at relatively small scale, and typically by NGOs. Moving forward, to reach the largest numbers of very poor households, either governments will have to implement the programs, or governments will fund implementation via subcontracts to local NGOs. Note that implementing the program at larger scale will mainly require increasing geographic coverage, rather than increasing the proportion of households reached in each village. This suggests that the smallish general equilibrium effects observed here are probably also representative of what one would expect from a larger program. Hence, the positive impacts generated by these programs are likely to be predictive of what a government could expect, if implemented similarly but at larger scale.

REFERENCES AND NOTES

1. World Bank, "The World Bank Annual Report 2013" (Washington, DC, 2013); https://openknowledge.worldbank.org/handle/10986/16091).
2. "A new global partnership: Eradicate poverty and transform economies through sustainable development" (United Nations High Level Panel, 2013); www.un.org/sg/management/pdf/HLP_P2015_Report.pdf.
3. S. D. Ashley, S. J. Holden, P. B. S. Bazeley, Livestock in Development (Organisation), *Livestock in Poverty-Focused Development* (Livestock in Development, Crewkerne, UK, 1999); www.theidlgroup.com/documents/IDLRedbook_000.pdf.
4. J. Dreze, Poverty in India and the IRDP Delusion. *Econ. Polit. Wkly.* **25** (29 September 1990); www.epw.in/review-agriculture/poverty-india-and-irdp-delusion.html.
5. R. V. Pulley, Making the poor creditworthy: A case study of the Integrated Rural Development Program. *World Bank Discuss. Pap.* **58** (1989); www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2000/06/28/000178830_98101903572627/Rendered/PDF/multi_page.pdf.
6. O. Bandiera, R. Burgess, S. Gulesci, I. Rasul, M. Sulaiman, Can entrepreneurship programs transform the economic lives of the poor? *LSE Work. Pap.* (2013); http://sticerd.lse.ac.uk/dps/eopp/eopp43.pdf.
7. A. U. Ahmed, M. Rabbani, M. Sulaiman, N. C. Das, The impact of asset transfer on livelihoods of the ultra poor in Bangladesh. *BRAC Res. Monogr. Ser.* **39** (2009); http://research.brac.net/new/component/k2/livelihoods.
8. M. S. Emran, V. Robano, S. C. Smith, Assessing the Frontiers of Ultrapoverty Reduction: Evidence from Challenging the Frontiers of Poverty Reduction/Targeting the Ultra-poor, an Innovative Program in Bangladesh. *Econ. Dev. Cult. Change* **62**, 339–380 (2014). doi: 10.1086/674110
9. N. C. Das, F. A. Misha, Addressing extreme poverty in a sustainable manner: Evidence from CFPR programme. *BRAC CFPR Work. Pap.* **19** (2010); http://research.brac.net/reports/cfpr_wp19.pdf.
10. In total, 10 sites were identified and programs implemented. Four are not included here for the following reasons: Yemen conducted a randomized evaluation, but has been delayed due to the civil conflict; a second India site, implemented by the microlender SKS, also conducted a randomized evaluation, but by a different set of researchers. It has not been included due to lack of comparability of data. They find no impact, due either to mistargeting individuals engaged in the labor markets, thus the grants generated substitution away from other income-generating activity; portfolio reallocation, in which productive asset grants were sold to pay down debt; or other data issues leading to lack of conclusive evidence (*21*). Two sites (Haiti, implemented by Fonkoze, and a second in West Bengal, India, implemented by Trickle-Up) did not employ experimental methods to measure their impact and are thus not reported here. Ford and CGAP also coordinated ethnographic research in several of the sites.
11. In Ghana, households received consumption support during the 6-month lean season for both years. They therefore received consumption support over the course of 2 years, but the total duration of support received was 14 months.
12. V. Alatas, A. Banerjee, R. Hanna, B. A. Olken, J. Tobias, Targeting the poor: Evidence from a field experiment in Indonesia. *Am. Econ. Rev.* **102**, 1206–1240 (2012). doi: 10.1257/aer.102.4.1206; pmid: 25197099
13. Grameen Foundation, PPI Construction; www.progressoutofpoverty.org/ppi-construction.
14. All exchange rates used in this paper are in PPP terms. We convert all monetary figures from local currency to USD PPP, at the year of the program's inception for cost data, and the year of the relevant survey for our results data. We then convert from USD PPP for that year to 2014 USD by multiplying by the ratio of the 2014 U.S. Consumer Price Index (CPI) to the U.S. CPI for the year in question. We use the following PPP rates: for Ethiopian Birr (ETB) in 2010, 2012, and 2013, 4.18, 6.45, and 6.66, respectively; for Ghanaian Cedis (GHS) in 2011, 2012, and 2014, 0.70, 0.79, and 0.91, respectively; for Honduran Lempiras (HNL) in 2009, 2012, and 2013, 9.77, 10.13, and 10.15, respectively; for Indian Rupees in 2007, 2009, and 2010, 11.76, 13.20, and 14.21, respectively; for Pakistan Rupees (PKR) in 2008, 2011, and 2013, 15.84, 24.35, and 26.83, respectively; for Peruvian Nuevo Soles (PEN) in 2011, 2013, and 2014, 1.48, 1.53, and 2.03, respectively. The U.S. CPIs used for 2007 to 2014 are, respectively, 207.3, 215.303, 214.537, 218.056, 224.939, 232.957, and 236.9111.
15. A. Deaton, O. Dupriez, Purchasing power parity exchange rates for the global poor. *Am. Econ. J. Appl. Econ.* **3**, 137–166 (2011). doi: 10.1257/app.3.2.137
16. J. B. Rotter, Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.* **80**, 1–28 (1966). doi: 10.1037/h0092976; pmid: 5340840
17. We present bounds for our treatment effects, depending on different assumption with respect to attrition, discussed in Results section.
18. In all countries, individuals were grouped into geographic block strata, which are included here as dummies for each block. In Honduras, Peru, and Ghana, rerandomization was performed to ensure balance on a set of variables. These variables are included as controls.
19. J. Kling, J. Liebman, L. Katz, Experimental analysis of neighborhood effects. *Econometrica* **75**, 83–119 (2007). doi: 10.1111/j.1468-0262.2007.00733.x
20. The ITT estimators can thus be interpreted as effect sizes relative to the control group.
21. J. Morduch, S. Ravi, J. Bauchet, "Failure vs. Displacement: Why an Innovative Anti-Poverty Program Showed No Net Impact" (CEI Working Paper Series 2012–05, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University, Tokyo, 2012); www.ier.hit-u.ac.jp/primced/documents/No32-dp_up_Pdf_2012_000.pdf.
22. M. Anderson, Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *J. Am. Stat. Assoc.* **103**, 1481–1495 (2008). doi: 10.1198/016214508000000841
23. The significance levels reported in Table 3 (* 10%, ** 5%, *** 1%) correspond to the naïve *p*-values, which can be inferred from the coefficient and standard errors.
24. Mental health questions were not asked in Pakistan in endline 1, whereas in India, women's empowerment was not asked about in endline 2, and so in both cases the correction is for only 9 outcome families when reporting country-specific indexed family outcomes in tables S3a to S3f . However, it is for 10 families in all other cases.
25. S. Subramanian, A. Deaton, The demand for food and calories. *J. Polit. Econ.* **104**, 133–162 (1996). doi: 10.1086/262020
26. This is gross livestock revenue, not income or net profit. On the expenditure side, it does not include fodder costs, which were not measured everywhere and were measured with considerable noise even where they were measured. On the profit side, it does not include unrealized capital gains (for example, as more calves are born, if they are not sold).
27. In some sites, individuals were shown a 10-rung ladder and asked, "How would you describe your satisfaction with life? If the top rung of this ladder (10) represents very satisfied and the lowest rung (1) represents very dissatisfied, where would you place yourself?" In others, individuals were shown five images of faces, and asked, "Which picture describes the current satisfaction level with your life, if the smiling face is the most satisfied and the crying/frowning face is the least satisfied?" In the latter case, the question was also scaled 1 to 10.
28. P. Brickman, D. T. Campbell, Hedonic relativism and planning the good society. *Adapt.-Level Theory*, 287–305 (1971).
29. J. Horowitz, C. Manski, Nonparametric analysis of randomized experiments with missing covariate and outcome data. *J. Am. Stat. Assoc.* **95**, 77–84 (2000). doi: 10.1080/01621459.2000.10473902
30. D. S. Lee, Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Rev. Econ. Stud.* **76**, 1071–1102 (2009). doi: 10.1111/j.1467-937X.2009.00536.x
31. Bandhan is also probably the organization that has the strongest links to BRAC, and the Bandhan program may have been run more similarly to BRAC's than the others. BRAC staff trained Bandhan staff at the onset of the program, for example.
32. The World Bank, "Staff guidance note on the application of the joint bank-fund debt sustainability framework for low-income countries" (82566, The World Bank, 2013), pp. 1–64.

33. C. Blattman, E. Green, J. Annan, J. Jamison, The returns to cash and microenterprise support among the ultra-poor: A field experiment. *Columbia Univ. Work. Pap.* (2014); http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2439488.
34. C. Blattman, N. Fiala, S. Martinez, Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *Q. J. Econ.* **129**, 697–752 (2014). doi: 10.1093/qje/qjt057
35. A. V. Banerjee, The two poverties. *Nord. J. Polit. Econ.* **26**, 129–141 (2000).
36. P. Dasgupta, D. Ray, Inequality as a determinant of malnutrition and unemployment: Theory. *Econ. J.* **96**, 1011–1034 (1986). doi: 10.2307/2233171
37. A. Banerjee, A. V. Banerjee, E. Duflo, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (PublicAffairs, New York, 2011).
38. O. Galor, J. Zeira, Income distribution and macroeconomics. *Rev. Econ. Stud.* **60**, 35–52 (1993). doi: 10.2307/2297811
39. A. Banerjee, A. Newman, Occupational choice and the process of development. *J. Polit. Econ.* **101**, 274–298 (1993). doi: 10.1086/261876
40. A. Banerjee, D. Karlan, J. Zinman, Six randomized evaluations of microcredit: Introduction and further steps. *Am. Econ. J. Appl. Econ.* **7**, 1–21 (2015). doi: 10.1257/app.20140287
41. M. Angelucci, D. Karlan, J. Zinman, Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *Am. Econ. J. Appl. Econ.* **7**, 151–182 (2015). doi: 10.1257/app.20130537
42. O. Attanasio, B. Augsburg, R. De Haas, E. Fitzsimons, H. Harmgart, The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *Am. Econ. J. Appl. Econ.* **7**, 90–122 (2015). doi: 10.1257/app.20130489
43. B. Augsburg, R. De Haas, H. Harmgart, C. Meghir, The impacts of microcredit: Evidence from Bosnia and Herzegovina. *Am. Econ. J. Appl. Econ.* **7**, 183–203 (2015). doi: 10.1257/app.20130272
44. A. Banerjee, E. Duflo, R. Glennerster, C. Kinnan, The miracle of microfinance? Evidence from a randomized evaluation. *Am. Econ. J. Appl. Econ.* **7**, 22–53 (2015). doi: 10.1257/app.20130533
45. B. Crépon, F. Devoto, E. Duflo, W. Pariente, Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *Am. Econ. J. Appl. Econ.* **7**, 123–150 (2015). doi: 10.1257/app.20130535
46. A. Tarozzi, J. Desai, K. Johnson, The impacts of microcredit: Evidence from Ethiopia. *Am. Econ. J. Appl. Econ.* **7**, 54–89 (2015). doi: 10.1257/app.20130475
47. A. Banerjee, S. Mullainathan, The shape of temptation: Implications for the economic lives of the poor. *Natl. Bur. Econ. Res.* **15973** (2010); www.nber.org/papers/w15973.
48. B. D. Bernheim, D. Ray, S. Yeltekin, Poverty and self-control. *Natl. Bur. Econ. Res.* **18742** (2013); www.nber.org/papers/w18742.
49. J. Haushofer, J. Shapiro, Welfare effects of unconditional cash transfers: Evidence from a randomized controlled trial in Kenya. *MIT Work. Pap.* (2013) http://www.princeton.edu/~joha/publications/Haushofer_Shapiro_UCT_2013.pdf).
50. S. de Mel, D. McKenzie, C. Woodruff, One-time transfers of cash or capital have long-lasting effects on microenterprises in Sri Lanka. *Science* **335**, 962–966 (2012). pmid: 22363007
51. D. McKenzie, C. Woodruff, What are we learning from business training and entrepreneurship evaluations around the developing world? *World Bank Policy Res. Work. Pap.* **6202** (2012).
52. Nor is the consumption increase simply the permanent income hypothesis in operation. If a household were capable of smoothing the income shock from the transfer perfectly, the increased long-term consumption would simply be the interest rate times the value of the assets transferred. The consumption increases we are observing are considerably higher than that.
53. M. Angelucci, G. De Giorgi, Indirect Effects of an Aid Program: How Do cash transfers affect ineligibles' consumption? *Am. Econ. Rev.* **99**, 486–508 (2009). doi: 10.1257/aer.99.1.486

ACKNOWLEDGMENTS

This study received approval from the Yale University Human Subjects Committee, IRB 0705002656, 1002006308, 1006007026, and 1011007628; the MIT Human Subjects Committee, IRB 0701002099; and from the Innovations for Poverty Action Human Subjects Committee, IRB Protocol 19.08January-002, 09December-003, 59.10June-002, and 10November-003.494.

Thanks to the Ford Foundation, 3ie, and U.S. Agency for International Development (USAID) [through the Financial Integration, Economic Leveraging, broad-based Dissemination and Support Leader with Associates (FIELD-Support LwA) managed by FHI 360] for funding. The contents of this paper are the responsibility of the authors and do not necessarily reflect the views of FHI 360, USAID, or the United States Government. Thanks to A. Agarwal, N. Barker, A. Kemmis Betty, C. Brewster, A. Bukari, D. Bullon Patton, S. De Marco, S. Devnani, M. Dieci, S. Fontenay, A. Garcia, Y. Guy, J. Manuel Hernández-Agramonte, S. Kant, S. Khan, H. Koizumi, M. Lowes, L. Luhana, J. Prasad Mukhopadhyay, E. Naah, M. Polansky, E. Safran, E. Salgado Chavez, D. Sánchez Liste, J. Severski, R. Strohm, H. Trachtman, and S. Vedder for outstanding research assistance and project management. The authors thank the leadership and staff at the implementing institutions for their partnership: Bandhan, Pakistan Poverty Alleviation Fund, Aga Khan Planning and Building Services Pakistan, Badin Rural Development Society, Indus Earth Trust, Sindh Agricultural and Forestry Workers Coordinating Organization, PLAN International Honduras, Organización de Desarollo Empresarial Feminino Social, Relief Society of Tigray, Presbyterian Agricultural Services (PAS), Asociación Arariwa, and PLAN International Peru. Thanks to F. DeGiovanni (Ford Foundation), S. Hashemi (BRAC University), and A. de Montesquiou and A. Latortue (CGAP) for their support and encouragement of the research. No authors have any real or apparent conflicts of interest, except that D.K. is on the Board of Directors of Innovations for Poverty Action, which participated in oversight of the implementation of the Ghana site. All data and code are available at the IPA Dataverse (doi: 10.7910/DVN/NHIXNT).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6236/1260799/suppl/DC1
Supplementary Text 1 to 4
Figs. S1 to S3
Tables S1 to S8

3 September 2014; accepted 6 April 2015
10.1126/science.1260799

Science
AAAS

A multifaceted program causes lasting progress for the very poor: Evidence from six countries
Abhijit Banerjee *et al.*
Science **348**, (2015);
DOI: 10.1126/science.1260799

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.