The Fading Treatment Effects of a Multi-Faceted Asset-Transfer Program in Ethiopia

Nathan Barker, Dean Karlan, Christopher Udry, Kelsey Wright[1]

We study the long-run effects of a big-push "graduation" program in Ethiopia in which very poor households received a one-time transfer of productive assets (mainly livestock), technical training, and access to savings accounts. After seven years, treatment effects on wealth and consumption remain economically meaningful but dissipated relative to the two- and three-year results. Treatment effects on other outcomes attenuated further. Based on absolute wellbeing (e.g., food security) not dropping, we argue that the treatment effect dissipation is driven primarily by improved living standards for control households, rather than losses of the previously accrued benefits for the treatment households.

JEL Classification: I32, I38, J24, O12

[1] Nathan Barker: University of Chicago, barkern@uchicago.edu; Dean Karlan: Northwestern University, CEPR, NBER, IPA, and J-PAL, karlan@ northwestern.edu; Christopher Udry: Northwestern University, CEPR, NBER, and J-PAL, christopher.udry@northwestern.edu; Kelsey Wright: Paris School of Economics, kelsey.wright@psemail.eu.
This study received approval from the Yale University Human Subjects Committee, IRB Protocol #1002006308, and from the Innovations for Poverty Action Human Subjects Committee, IRB Protocol #5.09December-003. The authors would like to thank the Relief Society of Tigray (REST) for partnering on the study, especially Teklewoini Assefa, Dr. Mulugeta Berhanu, Desta Gebremichael, and Zeratsion Fessha. Thanks to Sana Khan, Matthew Lowes, Hideto Koizumi, and Noor Sethi for outstanding research assistance. Thanks to Nathanael Goldberg for his expertise and work on project design and evaluation. JaRco Consulting assisted with the first wave of data collection, and Innovations for Poverty Action led subsequent waves. Teame Tesfay provided exceptional management of the follow-up surveys. 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. We also sincerely thank the study participants for their time.

I. Introduction

A central hypothesis within development economics is that poverty traps can explain why some individuals remain poor. At sufficiently low levels of wealth, individuals face a limited set of occupations, thus limiting wealth accumulation potential and in turn keeping them in low steady-state incomes and wealth levels (Galor and Zeira 1993; Banerjee and Newman 1993). Critically however, this literature hypothesizes that the same individuals could achieve a higher steady-state income if given a one-time infusion of capital. This theory, coupled with empirical evidence suggesting high short-run returns to capital in low-income countries (Udry and Anagol 2006; De Mel, McKenzie, and Woodruff 2008; Blattman et al. 2016; Blattman, Fiala, and Martinez 2014), has provided a basis for public policies that provide one-time capital transfers.

Consistent with this microeconomic theory and empirical results, several studies have found promising short- and medium-run evidence on the effectiveness of "graduation" programs at increasing the earnings and wealth of very low-income individuals in developing countries. In graduation programs, individuals receive a one-time transfer of productive assets (or cash to buy productive assets), coupled with training, consumption support, and improved access to savings; all together the aim is to push these individuals into a higher steady-state income. Randomized evaluations of graduation-style programs have found that a one-off transfer coupled with other initial support is sufficient to change the occupation individuals are engaged in, and cause a short-run increase in earnings and wealth (Bandiera et al. 2017; Banerjee et al. 2015; Bedoya et al. 2019; Brune et al. 2022). The theory inherent in these programs—that a one-off "big push" is sufficient to bring about higher steady state incomes—has been broadly adopted by policymakers. As of 2020, over 20 million households across 75 countries received some program centered around a one-off transfer (Andrews et al. 2021).

However, demonstrating the existence of poverty traps successfully overcome by these programs requires satisfying a more stringent set of criteria. In particular, if these programs push individuals over some critical threshold, we should expect to see the short-run gains of these programs persist, and moreover to see untreated households fail to clear the same threshold (Balboni et al. 2022). The available long-run evidence on graduation programs, while limited, is largely consistent with this hypothesis. In randomized evaluations of graduation programs in Bangladesh and India, treatment effects persist for at least seven and ten years, respectively (Bandiera et al. 2017; Banerjee, Duflo, and Sharma 2021). Showing that these programs lead to positive impacts over long time horizons across a broad variety of contexts is consistent with the hypothesis of the prevalence of poverty traps.

However, there are also theoretical reasons we might expect to see convergence in the longer run. If the threshold theory is right and a program fails to push individuals over a critical threshold, we should expect treated households to return to their pre-program equilibrium. Alternatively, if the constraints faced by the poor are not binding in the long-run, for example because individuals can borrow or save their way above a critical capital threshold, or because labor productivity in the occupations available to them is high, control households should eventually catch up to their treated peers (Buera, Kaboski, and Shin 2015).

While the "graduation" model cited above from India and Bangladesh found persistent long-run treatment effects (ten years), similar programs elsewhere often do not find long-run changes. For

example, a dissipation is observed in Uganda—cash transfers to young adults lead to income and wealth gains at two and four years, but the control group has caught up by year nine (Blattman, Fiala, and Martinez 2020). This contrasting result suggests additional value in further understanding the environments and program differences in which the impacts of capital transfer-centered program results are consistent with the presence of and escaping from a poverty trap.

When the long-run results are strong, a clearer case can be made for a poverty trap model. But when treatment effects fade, there are two broad interpretations. Is the treatment effect dissipation due to "seepage" in which treatment households slowly lose some of the accrued benefits they had experienced? Or due to control households catch-up? Naturally, this interpretation challenge is the logical consequence of properly constructing a counterfactual that is not merely the "pre" data for an individual. But using "pre" data for an individual as the counterfactual falls prey to countless omitted variable and selection biases, rendering the causal argument weak at best. However, while the trend-within-treatment may or may not be causal, we argue that it is still informative to simply ask whether there was any rescission of the absolute gains previously accrued to treatment households, which would be consistent with a poverty trap that was not surmounted by the treatment. If the decay of the treatment effect is accompanied instead by absolute gains by control households, some simple fixed-cost or production function non-convexity poverty trap models appear less useful.

We analyze the long-run effectiveness of a graduation program in Ethiopia, replicating the analysis protocols of Banerjee, Duflo, and Sharma (henceforth "BDS", 2021), who study a similar program in West Bengal, India. Both of these sites were part of a pooled-analysis of six-country graduation pilots, whose initial two and three year results were reported in Banerjee et al (2015). Each was a local adaptation of the NGO BRAC's graduation program.

In BDS, very poor households received either cattle, goats, or business inputs, and other support. The authors find strong treatment effects on consumption, income and revenues, and wealth at 18 months, three years, seven years, and ten years, with the size of treatment effects *increasing* between years three and seven. These accumulated gains appear to be driven in part by income diversification. Specifically, the graduation program relaxed liquidity constraints by year seven and ten that enabled members in treated households to migrate to more profitable locations; while control households also engaged in migration, it was to closer, less profitable cities and for shorter durations.

Our study consists of 925 individuals from the region of Tigray who were eligible for the Productive Safety Net Programme, Ethiopia's food-for-work program. Recruited subjects were enrolled in a lottery; winning households received a one-off transfer of either sheep and goats, oxen, bees, or inventory for petty trade, coupled with technical training and coaching, and were given access to local bank accounts (with the requirement that $250 must be saved before individuals could draw down their savings). The program yielded large and positive impacts on consumption, income and wealth after two and three years (Banerjee et al. 2015).

Our core result is that while there continue to be positive treatment effects on household welfare seven years after the initial program, long-run impacts have faded relative to the impacts present in earlier waves, in contrast to the growing treatment effects observed in India by BDS.

Two and three years after the initial transfer of assets, the graduation program led to large average increases in per capita consumption: treatment effects of 0.24 standard deviations at two years, 0.25 SDs at three years; and large increases in asset wealth: 0.95 SDs at both two and three years. Concurrently, we see large gains in the frequency with which individuals engage in livestock market transactions. Two years after the asset transfer, treated have bought an average of $74 more livestock in the last 12 months (relative to a control mean of $13—these and all other values are in 2021 USD Purchasing Power Parity (PPP) terms), and have sold $79 more livestock (relative to a control mean of $17). By year seven, treated households continue to have greater wealth and higher consumption, but the gap has closed meaningfully. The treatment effect on wealth has fallen to 0.43 SDs, and consumption to 0.13 SDs. The frequency of livestock transactions has also narrowed between the groups; in the seventh post-treatment year, treated households bought just $5 more livestock and sold $38 more.

Between years three and seven, control households caught up to treatment households in food security and livestock ownership, two domains where it is especially feasible to measure economic welfare in real terms. On our food security index, control households have improved by 0.71, 0.74 and 1.02 standard deviations in the two-, three- and seven-year surveys. For example, the share of control households reporting that all members get enough food everyday has risen from 34% at baseline to 83% at year seven. Similarly, we find (using a price index based on local goods, described in Appendix A) that livestock ownership among treated households has not declined, but rather held constant over time (from $2,308 at year-two to $2,469 at year-three to $2,449 at year-seven); the fading treatment effects reflect livestock accumulation on the part of control households.

Our results are thus consistent with the hypothesis that the graduation program accelerated treatment household consumption, income, and wealth growth, but not with the idea that it pushed households into an equilibrium they would have been unable to reach absent the program. Control households also experienced meaningful gains in living standards during the study period, engaged in many of the same economic behaviors (specifically livestock ownership and sales), and have closed the initial gap in earnings and wealth achieved by the program.

II. Experimental Design, Intervention and Data

The study took place in ten tabias (wards that include three to four villages) in Tigray Region, Ethiopia. Subjects were recruited from the government's food-for-work program, the Productive Safety Net Programme (PSNP). In our study area, PSNP participants earn chickpeas, teff and cooking oil in exchange for manual labor (e.g. digging trenches and engaging in land rehabilitation).

A local Community Task Force identified 1,000 PSNP participants for the study, on the basis of their having below average landholdings, low levels of livestock, high dependency ratios, and

limited external forms of income. Individuals also had to be capable of engaging in a physical activity, and not have an outstanding loan at any financial institution. After this initial screening, 925 individuals were contacted and deemed eligible.[2] Treatment status was determined via public lottery; we have no known instances of noncompliance with study design. Appendix Table A1 shows balance across the two groups at baseline. It is consistent with successful randomization.

The graduation program was implemented by Tigrigna NGO Relief Society of Tigray (REST), a large and well-known NGO in the region that offers a variety of services, including water and sanitation, livelihoods programs and education programs.

Treatment households were offered one of four productive asset options, each meant to be worth roughly $270 USD in value at the time ($1,371 2021 USD PPP): (1) 16 sheep or goats for fattening, (2) two oxen for fattening, (3) two beehives and colonies, and (4) a comparable value of inputs for petty trade. Most households chose some type of livestock: 62% chose shoats, 24% chose oxen, 10% beekeeping, and 4% petty trade. Each of the asset choices was accompanied by a specific technical training specific to their asset.

Households additionally received a bank account at the Dedebit Credit and Savings Institution (a microfinance institution) in the nearest market center. Individuals were required to accumulate savings equal to the value of the initial transfer prior to being able to withdraw their savings.[3] REST staff visited the program individuals for coaching on a regular basis over two years (initially weekly, then monthly). REST also offered quarterly refresher trainings of the core income generating principles. Unlike in some program contexts, treated households did not receive additional short-term cash or food transfers beyond what was already available to them (and to control households) via the PSNP program. Program activities ran from May 2010 to May 2012.

The program evaluated in BDS is also a local adaptation of the graduation program, with many similar components. As in Ethiopia, the most commonly chosen asset was goats in India (chosen by 52% of households), followed by cows (30%) and inputs for non-farm microenterprises (11%). Individuals received weekly cash transfers of INR 90 ($10.03) for 13-40 weeks (depending on the asset chosen), received coaching on a regular basis, and were required to save INR 10 ($1.11) per week when meeting with the implementing staff.

The research NGO Innovations for Poverty Action conducted four waves of surveys. The baseline survey took place two months prior to the lottery. Follow-up surveys took place at the two-, three-, and seven-year marks. In each, we administered both a household survey, and up to two adult surveys. The household survey included household member characteristics, housing, assets, income, consumption, food security, health spending and shocks. The adult survey was administered to the household head (and their spouse, if married); questions included time use,

[2] 19 were not found; 52 were found but deemed ineligible on the basis of existing loans, and 4 were discovered to be duplicates at the time of surveying.

[3] This was done because of the Government of Ethiopia's general opposition to "handouts;" this program feature ensured that the program was "like a loan," though unlike a loan, households did not have to repay anyone, but rather to save the amount prior to withdrawing the funds.

control over decision-making, physical and mental health, and participation in political activities. A complete timeline of the program and surveying can be found in Appendix Table A2.

Attrition is low. Of the 925 sample households, we successfully interviewed 99% in our two-year survey, 98% in our three-year survey, and 96% in our seven-year survey. Appendix Table A3 shows that attrition does not systematically differ between treatment and control households.

III. Results

Empirical Strategy

We estimate average treatment effects of the graduation program on outcome *y* for household *h* at time *t* with the specification:

$$y_{ht} = \alpha + \beta\, Treat_h + \delta\, y_{ho} + \mu_v + \epsilon_{ht}$$

Where $Treat_h$ is an indicator variable for whether the household was randomly chosen to receive the program, y_{ho} is the outcome variable at baseline (coded as 0 at missing if baseline, with an additional indicator variable for baseline missingness), μ_v are village fixed-effects, and ϵ_{ht} is our error term. Standard errors are Huber-White heteroskedastic for outcomes reported at the household level. For outcomes reported at the adult-level (i.e. with 1-2 members surveyed per household), we cluster our standard errors at the household-level.

Indexed Family Outcomes

To provide an aggregate view of how the graduation program affects household and individual welfare over time, we report results on indexed family outcomes, following the grouping of families laid out in Banerjee et al. (2015) and in BDS. We follow Kling, Liebman, and Katz (2007) and create z-score indices for each family of outcomes. We standardize against the baseline whenever possible, or against the control group at year two (when any relevant variables were not collected at baseline).[4] Additionally, we correct for multiple hypothesis testing using the Benjamini-Hochberg step-up procedure to control for the false discovery rate (FDR) (Benjamini and Hochberg 1995), following the procedure discussed in Anderson (2008).

Results are reported in Table 1 for our three follow-up surveys conducted at the two-, three- and seven-year marks, reported in Columns (1), (2) and (3), respectively. To facilitate comparison with the results realized in India, we report the results from India at 1.5 years, three years, and seven years in Columns (4), (5), and (6) (these results come directly from Table 1 of BDS, and are inflated to be in 2021 USD PPP terms). Following BDS, we report outcomes on political and women's empowerment in Appendix Table A4, because of space constraints.[5]

[4] Three measures—per capita consumption, asset wealth and productive time use—are aggregates. In these cases, we index against the aggregate rather than the mean of the component parts.

[5] We still adjust our q-values for these families; i.e. we adjust for the fact we are testing ten indexed family outcomes.

Three key patterns emerge from Table 1.

First, the evaluation took place during a period of meaningfully improving living standards for control households. Asset ownership among control households has improved by 1.16 standard deviations (SDs) at the seven-year mark relative to baseline, and similarly, food security has improved by 1.02 SDs. Relative to the two-year mark, real per capita consumption has increased by 0.49 SDs at the seven-year mark, and income and revenues by 1.25 SDs. These changes in control means suggest that the environment in which the program took place was not one in which households were stagnating in a low-level equilibrium, but rather an environment of solid economic growth. Indeed, from 2011-2017, GDP grew in Ethiopia at an average rate of 10% per year—leading to GDP per capita more than doubling (117% increase) during the study period.

Second, treatment effects on economic outcomes (wealth, consumption, food security, income and revenues, financial inclusion, productive time use) all follow a similar pattern: large and positive effects at years two and three, and with positive coefficients, albeit of reduced magnitude by year seven (and in many cases, a loss of statistical significance). For example, the estimated treatment effect on asset ownership has fallen from 0.95 SDs at years two and three to 0.43 SDs at year seven. Treatment effects on income and revenues and financial inclusion both exceeded 1 SD at their peak at year two (1.41 SDs in the case of income and revenues, 1.85 in the case of financial inclusion). By year seven, both point estimates are still moderately sized (0.24 and 0.30 SDs, respectively), but in neither case can we reject the null of no effect. Of note, because variance in our sample is growing over time, we are less powered to detect effects of a given size by the year-seven mark.

Our consumption estimates paint an intermediate picture: the decline from year-three to year-seven is relatively modest: 0.25 (p-value<0.01) to 0.17 SDs (p-value=0.036). However, given the large number of families we are testing (and the lower rate of statistical significance across the ten families at year seven, leading to a greater FDR adjustment), the adjusted q-value on per capita consumption is 0.196 at year-seven, suggesting ambiguity about the persistence of this effect.

Third, we observe very limited evidence of treatment effects on downstream outcomes. In no wave can we reject the null that the program had no effect on mental or physical health, or women's decision-making (the last of which is reported in Appendix Table A4, given space constraints). This result stands in contrast to BDS, which finds at 1.5 and seven years that the program led to improved perceived physical and mental health.

Consumption and Food Security

In Table 2, we document the evolution of treatment effects on consumption and food security in Ethiopia and India, with a focus on the extent to which the program pushes individuals out of extreme poverty.

In Ethiopia, control household living standards are improving considerably in the time horizon studied. At year seven, control household per capita consumption is 26% higher than the baseline sample mean. The share of households getting enough food everyday has grown from 34% to

83%. Treatment effects are initially positive on both consumption and food security—by the three-year mark, per capita consumption is 18% among treatment households; food security has improved by 0.15 SDs. By the seven-year mark, consumption gains fade but persist (9% higher), but we can no longer reject that treatment effects are zero on our measures of food security.

While these accumulated gains are sizable, treated households are still very poor by most standards. By the seven-year mark, control household per capita consumption is equal to $2.23 per day, and treatment consumption $2.43. Mean treatment household consumption has thus just exceeded the World Bank's extreme poverty threshold of $2.38 per day.[6]

The treatment effects in India initially follow a similar initial trajectory to Ethiopia, with strong treatment effects in years two and three. In contrast, by year seven, consumption and food security treatment effects have grown in India. This divergence cannot be explained by the fact that control household living standards have improved in Ethiopia—in India, control households have also experienced large consumption growth, of 64% between baseline and year seven. In fact, BDS hypothesize that the initial positive program effects might have enabled treatment households to better take advantage of the strong macroeconomic conditions.

Income and Revenues

The same pattern of divergence between Ethiopia and India is also present in Table 3, in which we examine income and revenues, (further disaggregated for Ethiopia in Table 4).

In Ethiopia, we observe strong and statistically significant treatment effects on income and revenues by years two and three, with results fading somewhat by year seven. On livestock revenues, we see extremely large effects at the two-year mark; treated individuals have $128 more per month (relative to a control mean of $32). This effect fades by years three and seven to $31 and $40 per month, respectively. At the seven-year mark, these declining results appear to be driven in part by control households increasing the extent to which they engage in livestock market transactions—their mean increases from $30 per month at year three to $90 per month at year seven. In India, treated households also experience gains in livestock revenues at the two-year mark. Again, in contrast to Ethiopia, these treatment effects have increased by the seven-year mark.

In both Ethiopia and India, treatment households diversify their income sources beyond the activities directly promoted by the program. In Ethiopia, this is primarily through increased agricultural activity. Net agricultural profits increase at years two and three (and there are positive point estimates at year seven, though we cannot reject the null of no effect). As shown in Table 4, in all three follow-up waves the total household expenditure on agriculture increases, as does the land cultivated. In years two, three and seven, the acres cultivated by the household increases by 0.15, 0.22, and 0.25, respectively (representing 11%, 14%, and 18% increases). In years two and three, these increases translate to increases in agricultural revenue; by year-seven we can no longer reject that treatment and control means are equal.

[6] Inflation adjusted

In India, the degree of income diversification is more multifaceted. In addition to experiencing gains in livestock, the asset promoted by the program, treated households in India have increased income via fishing and horticulture, and via non-farm enterprise profits. Moreover, as in other outcomes, these treatment effects have increased from year-three to year-seven.

Another important difference between India and Ethiopia is the degree to which migration and remittance are an important and growing part of sample households' earnings portfolios. In neither country are there statistically significant differences between treatment and control households at the seven-year mark. However, in India, both treatment and control household remittances have grown substantially from years three to seven—control household remittances have grown by 171% in this time period. By year ten, there is a statistically significant difference (remittances are 52% higher). This effect is driven by household members spending more time away, being less likely to migrate to Kolkata (the nearest urban center) and earning more when they do migrate. BDS hypothesize that this was a key pathway to growing treatment effects in India. As the macroeconomic conditions improved, treatment households had more liquidity, and could thus finance more profitable migration spells.

While we only observe effects through year seven in Ethiopia, we do not observe evidence suggesting a shift towards migration and remittances. Mean control group remittances grow from years two to three but are more or less constant between years three and seven (control mean of $9.06 per month at year three, $10.34 by year seven). Moreover, when we examine income channels further in Table 4, we do not see evidence of growth on the extensive margin in households receiving any remittances (for treatment households, the share receiving any remittances at years three and seven are 16% and 18%, respectively, while for control households, the equivalent values are 24% and 14%, respectively).

One plausible difference between the two settings is the degree to which non-financial constraints to migration were binding. The study in Ethiopia took place during a period of political instability and ethnic tensions, which could have discouraged migration outside of Tigray, to parts of the country where the Tigrigna are an ethnic minority. Since 1991, a Tigrigna-dominated coalition had governed Ethiopia. Mass protests and government crackdowns outside of Tigray began in 2014 (year four), leading to the end of this government in 2018, a year after our seven-year survey. Consistent with this, we observe low rates of individuals leaving Tigray for another part of Ethiopia—just 1% by year three, and 5% by year seven. In fact, households were more likely to have a member who left Tigray *for another country* (8% and 9% at years three and seven, respectively) than for another region of Ethiopia. On neither of these margins do results differ by treatment status.

Cost-Benefit Analysis

While the benefits of the graduation program to treatment households were sizable, intervention costs are also substantial: in 2021 USD PPP terms, the total cost of the program was $4,011 per participant, reported in Panel A of Appendix Table A5. Of this, $1,371 (34%) constituted a direct transfer to the household in the form of livestock or business inputs; the other two-thirds included a combination of implementing staff salaries, training, supervision costs, and indirect

expenses. Notably, the expenditures were more than triple that of the per-person expenditures in India, though of the six sites reported in Banerjee et al. (2015), Ethiopia is the third-cheapest. While in practice, some of the costs were incurred over the two years the training and coaching continued, we assume that all costs were incurred immediately at time zero when discounting.

We calculate benefits on the basis of per capita consumption, reported in Panel B of Appendix Table A5. In particular, we assume that the full benefit to households is equal to the sum of all consumption treatment effects. We make three key additional assumptions. First, we assume that the impacts on consumption at the one-year mark are equal to those at the two-year mark. Second, we assume that the decline in consumption treatment effects from year-three to year-seven followed a linear decline. Third, we assume that this linear rate of decline from year-three to -seven continued/continues, up to the final year at which treatment effects would be greater than zero under this decline (year 16, in 2026). We annualize these consumption benefits, and apply a 5% annual discount rate, following the social discount rate used by the World Bank and International Monetary Fund (International Monetary Fund 2013). Our estimated benefits are restricted to consumption effects. We are therefore neglecting any insurance benefit associated with increased wealth (including via the directly transferred assets), nor are we incorporating any non-pecuniary benefits, for example, through feeling more politically empowered.

Under these assumptions, we find that the total costs and total benefits are very similar—total benefits to the household equal 97% of costs. This result depends fundamentally on the fact that while program costs are incurred immediately, benefits are gradual and extend over a long-time horizon. In total (as shown in Figure 1, and Panel C of Appendix Table A5), with no discount rate applied, the total projected benefits as more than double the program costs. Benefits are equal to costs at a social discount rate of 4.4%.

We quantify the uncertainty of our benefit-cost ratio estimates using bootstrap simulations. For each of 10,000 repetitions, we construct a bootstrap sample, re-estimate consumption treatment effects on this sample, and calculate our benefit-cost ratio on these estimated values. (This procedure is further described in Appendix B). We find that benefits exceed costs in 52% of our simulations; our 95% confidence interval of the benefit cost-ratio is [0.42, 2.36]. These results thus suggest sizable uncertainty about the relative ratio of costs and benefits, emphasizing the further importance of accumulating additional evidence on the effectiveness and costs of these programs.

Moreover, the results in India are important because they suggest the possibility of large treatment effects with a much cheaper version of the program. In India, the direct transfers to households equaled roughly half of total program costs, compared to roughly a third in Ethiopia. Work on a graduation program in Ghana has found that neither simply providing access to savings, nor providing a one-off transfer of goats with no additional training, is sufficient to achieve the same impacts as those observed with the full graduation program (Banerjee et al. 2022), suggesting that the solution is not as straightforward as simply eliminating core components. We view the question of how to either pare down the cost of these programs (perhaps through reduced costs associated with scale up), or how to increase the effectiveness (e.g., by strengthening the socioemotional skills of individuals receiving the assets, as in Barker et al. (2022) and Bossuroy et al. (2022)), as important areas for future study.

IV. Conclusion

In Ethiopia, a transfer of productive assets, coupled with training, coaching and access to savings accounts induced large treatment effects on consumption, income and wealth two and three years later, and led to positive but diminished treatment effects seven years after the initial transfer. These fading treatment effects took place in a context in which control household living standards were improving consistently. We do not find evidence of dissipation of any of the previously accrued absolute improvements for treatment households. Consumption and related welfare measures continued to grow for both treatment and control households throughout the study period, with the gap across groups closing over time. These results differ from those of a similar program in India.

A potentially important difference between the results in India and the ones here are the degree and nature of income diversification that followed from the initial, positive treatment effects on income. In both contexts, in the short-run, treatment households experienced gains in livestock revenues. In Ethiopia, households also engaged in income diversification, albeit narrowly in the domain of agriculture. Households cultivate more land, invest more on their farms, and in the short run, have higher revenues, although by year seven this effect is no longer significant. In India, the degree of income diversification was more sizable—households increased their earnings from fishing and horticulture, and from non-farm enterprises.

Most importantly however, the two programs differ in the extent to which they facilitated additional migration. In India, both treated and control households increase their rate of migration and remittances by year seven; by year ten, treated households have 52% more remittances than control households. This appears to be driven by households being more likely to migrate to more distant locations, and for longer periods of time. In Ethiopia, while our results only extend through year seven, at no point do we observe an increase in remittances in Ethiopia, and migration rates to parts of Ethiopia outside of Tigray remained low for both treated and control households. These results suggest that either domestic migration to other parts of Ethiopia outside of Tigray were deemed unappealing, and/or that the constraints to such migration remained binding.

Our results are largely inconsistent with the hypothesis that the program unlocked a poverty trap for poor households in Ethiopia, in contrast to what has been observed in India and Bangladesh (in which results are consistent with, albeit perhaps not dispositive of, the hypothesis that the programs pushed individuals over some critical threshold). This difference also matters from a policy perspective, as NGOs and governments consider the scale-up of graduation programs.

In a public finance sense, the question is whether from a cost-effectiveness perspective they should be compared to short-run fixes of humanitarian poverty-based crises or compared to attempts to address long-run income and inclusive market development approaches. In a practical sense, this has implications for the number of years over which one would reasonably project benefits in a cost-benefit analysis. Of the six graduation pilots studied in Banerjee et al. (2015), for none did the discounted benefits from the first three years exceed the costs. The five

that did have benefits in excess of costs only did so with assumptions that the benefits continued past the three years, which seemed reasonable to assume some continuation given the similar results at two and three years. The predicted continued benefits have been realized in India, but have been positive but smaller in Ethiopia. The fading of results here, compared to the persistence elsewhere, suggest continued long-run study of these programs is essential in order to understand how and when such results persist.

Coda: Civil Conflict in Tigray

As in any empirical study, the effects of programs or actions depend on how they interact with the external conditions, including both positive and negative shocks (Rosenzweig and Udry 2020). Unfortunately, in the time since our most recent data collection, the area where our study took place has experienced a severe, negative shock in the form of violent conflict and displacement.

In the time since the end of the Civil War in 1992 until 2018, the Tigray People's Liberation Front (TPLF), the party sharing an ethnicity with our sample, has been a central power in Ethiopia's ruling coalition. This changed in 2018, when a new ruling coalition was formed that excluded the TPLF from power and led to tensions between the TPLF and the newly ruling Prosperity Party. These tensions transitioned to violent conflict in late 2020, with the proximate cause of disputes over the legitimacy of Tigrayan elections held during the COVID pandemic. As of October 2022, this conflict has led to an estimated 385,000-600,000 deaths and more than 2 million displacements (York 2022); reporting suggests the study area was heavily affected.

As of November 2022, a peace agreement was signed between the two parties; while conflict had not fully ceased, the parties agreed to a coordinated disarmament.

Thus, the extent to which treatment effects persisted or faded in our study site will depend on how the program interacted with this violent conflict and its immediate resolution, rather than with a context of macroeconomic growth and stability, as was the case from baseline through our three follow-ups. Previous work finds positive impacts of the graduation program in the face of conflict (Bedoya et al. 2019 and Brune et al. 2022 report strong impacts in Afghanistan and modestly positive impacts in Yemen, respectively). Clearly, meaningful uncertainty remains about the relationship between graduation programs and resilience, and how treatment effects will evolve for participants in this program in Tigray, Ethiopia.

References

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Andrews, Colin, Aude de Montesquiou, Inés Arévalo Sánchez, Puja Vasudeva Dutta, Sadna Samaranayake, Janet Heisey, Timothy Clay, and Sarang Chaudhary. 2021. *The State of Economic Inclusion Report 2021: The Potential to Scale*. World Bank Publications.

Balboni, Clare, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil. 2022. "Why Do People Stay Poor?" *The Quarterly Journal of Economics* 137 (2): 785–844.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2017. "Labor Markets and Poverty in Village Economies." *The Quarterly Journal of Economics* 132 (2): 811–70.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236).

Banerjee, Abhijit, Esther Duflo, and Garima Sharma. 2021. "Long-Term Effects of the Targeting the Ultra Poor Program." *American Economic Review: Insights* 3 (4): 471–86. https://doi.org/10.1257/aeri.20200667.

Banerjee, Abhijit, Dean Karlan, Robert Osei, Hannah Trachtman, and Christopher Udry. 2022. "Unpacking a Multi-Faceted Program to Build Sustainable Income for the Very Poor." *Journal of Development Economics* 155: 102781.

Banerjee, Abhijit V, and Andrew F Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101 (2): 274–98.

Barker, Nathan, Gharad Bryan, Dean Karlan, Angela Ofori-Atta, and Christopher Udry. 2022. "Cognitive Behavioral Therapy among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress." *American Economic Review: Insights*.

Bedoya, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro. 2019. "No Household Left behind: Afghanistan Targeting the Ultra Poor Impact Evaluation." National Bureau of Economic Research.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289–300.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *The Quarterly Journal of Economics* 129 (2): 697–752.

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

Blattman, Christopher, Eric P Green, Julian Jamison, M Christian Lehmann, and Jeannie Annan. 2016. "The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda." *American Economic Journal: Applied Economics* 8 (2): 35–64.

Bossuroy, Thomas, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C. Thomas, Christopher Udry, and Julia Vaillant. 2022. "Tackling Psychosocial and Capital Constraints to Alleviate Poverty." *Nature* 605 (7909): 291–97.

Brune, Lasse, Dean Karlan, Sikandra Kurdi, and Christopher Udry. 2022. "Social Protection amidst Social Upheaval: Examining the Impact of a Multi-Faceted Program for Ultra-Poor Households in Yemen." *Journal of Development Economics* 155: 102780.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2015. "Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective." *Economics* 7 (1): 409–36.

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. “Returns to Capital in Microenterprises: Evidence from a Field Experiment.” *The Quarterly Journal of Economics* 123 (4): 1329–72.

Galor, Oded, and Joseph Zeira. 1993. “Income Distribution and Macroeconomics.” *The Review of Economic Studies* 60 (1): 35–52.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica* 75 (1): 83–119.

Rosenzweig, Mark R., and Christopher Udry. 2020. “External Validity in a Stochastic World: Evidence from Low-Income Countries.” *The Review of Economic Studies* 87 (1): 343–81.

Udry, Christopher, and Santosh Anagol. 2006. “The Return to Capital in Ghana.” *American Economic Review* 96 (2): 388–93.

York, Geoffrey. 2022. “Surge of Dehumanizing Hate Speech Points to Mounting Risk of Mass Atrocities in Northern Ethiopia, Experts Say.” *The Globe and Mail*, October 21, 2022. https://www.theglobeandmail.com/world/article-surge-of-dehumanizing-hate-speech-points-to-mounting-risk-of-mass/.

TABLES

Table 1: Indexed Family Outcomes

	Ethiopia			India		
	(1)	(2)	(3)	(4)	(5)	(6)
	2 Years	3 Years	7 Years	1.5 Years	3 Years	7 Years
Asset Ownership	0.950***	0.947***	0.430***	0.217*	0.389***	0.814***
	(0.09)	(0.09)	(0.11)	(0.11)	(0.10)	(0.13)
q-value	0.001	0.001	0.001	0.041	0.001	0.001
Control Mean	0.374	0.519	1.159	-0.2	-0.25	-0.46
Baseline Mean		0.000			0.000	
Per Capita Consumption	0.239***	0.250***	0.172**	0.311***	0.292***	0.717***
	(0.07)	(0.05)	(0.08)	(0.08)	(0.08)	(0.13)
q-value	0.001	0.001	0.196	0.00	0.00	0.00
Control Mean	0.000	-0.0948	0.487	0.35	0.85	1.09
Baseline Mean		-			0.00	
Food Security Index	0.116**	0.145***	0.0351	0.184***	0.251***	0.431***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.06)	(0.06)
q-value	0.022	0.003	0.464	0.00	0.00	0.00
Control Mean	0.706	0.742	1.024	0.35	0.94	1.09
Baseline Mean		0.000			0.00	
Income and Revenues Index	1.411***	0.411***	0.244	0.145**	0.172***	0.334***
	(0.16)	(0.12)	(0.24)	(0.08)	(0.07)	(0.07)
q-value	0.001	0.001	0.464	0.04	0.02	0.00
Control Mean	0.000	0.512	1.247	0	0	0
Baseline Mean		-			-	
Financial Inclusion Index	1.853***	0.779***	0.301	-0.004	0.192***	0.181
	(0.12)	(0.13)	(0.25)	0.04	(0.06)	(0.14)
q-value	0.001	0.001	0.449	0.26	0.00	0.05
Control Mean	0.000	0.0473	0.736	0.14	0.3	0.67
Baseline Mean		-			0.000	
Physical Health Index	-0.0158	0.0499	-0.116	0.061***	0.027	0.13***
	(0.05)	(0.04)	(0.08)	(0.03)	(0.03)	(0.03)
q-value	0.293	0.104	0.340	0.03	0.16	0.00
Control Mean	-0.0749	0.0112	-0.0525	0.913	0.921	0.987
Baseline Mean		0.005			0.000	
Mental Health Index	0.0527	-0.0384	-0.0872	0.115***	0.012	0.0113***
	(0.06)	(0.06)	(0.06)	(0.03)	(0.04)	(0.02)
q-value	0.148	0.210	0.361	0.00	0.33	0.00
Control Mean	0.00	0.00	0.00	0.32	0.75	1.09
Baseline Mean		-			0.000	
Productive Time Use	0.278***	0.224***	0.0805	0.285***	0.102**	0.165***
	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
q-value	0.001	0.001	0.340	0.001	0.018	0.000
Control Mean	0.508	0.604	0.463	0.23	0.28	-0.04
Baseline Mean		0.002			0.000	

Table 1 reports results from intention-to-treat effects of the graduation program on indexed family outcomes. Columns (1) to (3) report results from Ethiopia, and Columns (4) to (6) from India, as presented in Banerjee et al. 2021. Each cell reports results from a separate regression. Observations range from 889 to 915 for the household level outcomes in Ethiopia and from 679 to 875 in India, and from 723 to 1307 for the adult-level outcomes in Ethiopia and from 1,229 to 1,950 in India. For asset ownership, consumption and time use, outcomes are aggregated and then rescaled; other outcomes are a z-score index of outcomes variables in the given category. Outcomes are either (a) standardized so the baseline has mean zero and standard deviation one (done whenever possible), or (b) standardized so the endline 1 control group has mean zero and standard deviation one (done in cases where we do not have baseline data, or the baseline components differed from the components in subsequent waves). Reported q-values are sharpened using the false discovery rate procedure detailed in Anderson (2008). They reflect a correction for 10 family outcomes in Ethiopia, and 9 family outcomes in India. Only 8 are presented here given space constraints, the full set for Ethiopia is reported in Appendix Table A3. The full list of variables used to construct each index is reported in Appendix FIXME. Each regression controls for baseline values, and for village-level strata. Standard

Table 2: Monthly Consumption and Food Security

	Ethiopia			India		
	(1)	(2)	(3)	(4)	(5)	(6)
	2 Years	3 Years	7 Years	1.5 Years	3 Years	7 Years
Per Capita Consumption	8.210***	8.561***	5.896**	8.352***	7.83***	19.222***
	(2.32)	(1.83)	(2.81)	(2.031)	(2.115)	(3.350)
Control Mean	50.16	46.91	66.88	54.45	67.86	74.25
Baseline Mean		53.04			45.05	
Per Capita Food Consumption	2.908	2.345**	1.082	5.929***	3.340***	10.811***
	(1.78)	(1.00)	(1.42)	(1.321)	(1.273)	(1.914)
Control Mean	32.59	30.02	38.34	36.80	41.51	41.75
Baseline Mean		36.974			29.59	
Per Capita Nonfood Consumption	4.413***	4.610***	3.857**	2.358***	4.43***	8.339***
	(1.02)	(1.11)	(1.84)	(1.196)	(1.259)	(1.859)
Control Mean	16.15	16.2	26.42	17.66	26.35	32.51
Baseline Mean		16.066			15.48	
Per capita durable good consumption	0.497	1.206***	0.387	-0.394	0.974***	2.732***
	(0.40)	(0.34)	(0.52)	(0.414)	(0.426)	(0.561)
Control Mean	1.422	0.685	2.116	2.52	2.13	2.51
Baseline Mean		-			1.07	
Everyone in HH gets enough food everyday	0.0386	0.0766**	0.0234	0.074***	0.141***	0.205***
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Control Mean	0.64	0.64	0.826	0.11	0.42	0.59
Baseline Mean		0.335			0.11	
No one HH went whole day without food	0.0293*	0.0411***	-0.00719	0.128***	0.038*	0.095***
	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)	(0.02)
Control Mean	0.913	0.921	0.987	0.68	0.85	0.83
Baseline Mean		0.745			0.28	
No Children Skipped Meals	0.0447**	0.0468**	0.0113	0.032	0.085***	0.045*
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Control Mean	0.845	0.833	0.926	0.75	0.86	0.87
Baseline Mean		0.489			0.51	
Share of total consumption on non-grain items	0.0384***	0.0366***	0.0129*	0.00634	0.0211***	0.0240***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Control Mean	0.612	0.58	0.686	0.708	0.714	0.773
Baseline Mean		0.525			0.237	

Table 2 reports results from intention-to-treat effects of the graduation program on consumption and food security-related outcomes. Columns (1) to (3) report results from Ethiopia, and Columns (4) to (6) from India, as presented in Banerjee et al. 2021. Each cell reports results from a separate regression. Observations range from 889 to 915 for Ethiopia and from 679 to 875 in India. All outcomes in Ethiopia are reported in 2021 USD in PPP terms. Each regression controls for baseline values (when available), and for village-level strata. Standard errors are Huber-White heteroskedastic. Additional food security outcomes are reported in Appendix Table A4.

Table 3: Monthly Income and Revenue

	Ethiopia			India		
	(1)	(2)	(3)	(4)	(5)	(6)
	2 Years	3 Years	7 Years	1.5 Years	3 Years	7 Years
Livestock Revenues	128.5***	30.85***	40.20*	11.342***	8.495	30.143***
	(8.71)	(6.25)	(20.98)	(2.589)	(6.918)	(5.703)
Control Mean	32.48	29.51	89.9	3.68	8.83	10.73
Baseline Mean		8.437			0.00	
Agricultural Profits (ETH) / Fishing & Horticulture (IND)	5.046**	6.849***	4.296	20.637***	34.339***	119.81***
	(2.30)	(2.38)	(3.49)	(6.408)	(7.626)	(16.750)
Control Mean	25.72	34.62	33.46	50.96	66.89	114.05
Baseline Mean		16.554			18.02	
Non-Farm Enterprise Profits	12.03**	-0.856	-1.346	8.765*	27.770***	74.735***
	(4.91)	(3.13)	(5.78)	(5.018)	(6.918)	(15.771)
Control Mean	4.074	16.300	7.422	40.07	54.70	100.05
Baseline Mean		2.3			14.64	
Income from Wage Labor (including workfare	-2.797	-2.289	-2.984	5.62	5.117	98.431***
	(2.95)	(2.68)	(6.76)	(9.382)	(15.119)	(28.449)
Control Mean	34.78	29.83	89.97	117.70	240.91	333.15
Baseline Mean		9.23			-	
Rating of Economic Status (1/10)	1.126***	0.901***	0.0062	0.204***	0.297***	1.575***
	(0.15)	(0.12)	(0.13)	(0.07)	(0.08)	(0.14)
Control Mean	3.732	4.335	4.989	2.77	3.36	4.73
Baseline Mean		3.742			1.97	
Productive Asset Value, Indexed	0.776***	0.761***	0.374***	0.444***	0.571***	0.795***
	(0.08)	(0.07)	(0.10)	(0.09)	(0.07)	(0.08)
Control Mean	0.0685	0.173	0.863	-0.23	-0.3	-0.4
Baseline Mean		0.000			0.000	
Remittances Received	1.554	1.214	4.272	-	4.087	9.808
	(1.18)	(2.89)	(3.90)	-	(2.619)	(7.137)
Control Mean	2.751	9.057	10.34	-	14.241	38.555
Baseline Mean		-			-	

Table 3 reports results from intention-to-treat effects of the graduation program on income and revenue outcomes. Columns (1) to (3) report results from Ethiopia, and Columns (4) to (6) from India, as presented in Banerjee et al. 2021. Each cell reports results from a separate regression. Observations range from 889 to 915 for Ethiopia and from 679 to 875 in India. All outcomes in Ethiopia are reported in 2021 USD in PPP terms. Each regression controls for baseline values (when available), and for village-level strata. Standard errors are Huber-White heteroskedastic.

Table 4: Income Mechanisms, Ethiopia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total Value of Livestock	Total Value of Livestock Sold per Month, last 12 months	Total Value of Livestock Bought per Month, last 12 months	Revenue from Agriculture	Expenditure on Agriculture	Total Acres Cultivated, Major Growing Season	Received any remittances	Earned any wage income (including from workfare program)
Treatment (ITT): Two Year	967.3***	79.24***	73.67***	7.997***	4.032***	0.153**	-0.029	0.004
	(91.230)	(5.487)	(5.275)	(2.552)	(1.166)	(0.068)	(0.023)	(0.007)
Observations	914	914	914	915	915	915	911	915
Control mean	1341	17.66	12.97	37.49	11.78	1.44	0.15	0.99
Treatment (ITT): Three Year	988.1***	30.33***	30.91***	8.356***	2.911***	0.221***	-0.0802***	-0.0371**
	(88.980)	(5.520)	(5.334)	(2.508)	(0.807)	(0.066)	(0.026)	(0.017)
Observations	908	908	908	908	908	908	908	915
Control mean	1481	13.86	10.76	41.97	7.35	1.56	0.24	0.94
Treatment (ITT): Seven Year	419.2***	38.18*	5.353*	6.428	3.690***	0.254**	0.038	-0.0816***
	(108.100)	(19.990)	(3.151)	(4.040)	(1.309)	(0.098)	(0.025)	(0.022)
Observations	889	889	889	889	889	889	889	889
Control mean	2030	67.06	22.95	54.15	20.69	1.43	0.14	0.91
Baseline Mean	1107.44	-	-	81.37	14.01	0.86	-	0.11

Table 4 reports results from intention-to-treat effects of the graduation program on mechanisms through which household-level income and revenues have evolved. All financial outcomes are reported in 2021 USD in PPP terms. Each regression controls for baseline values (when available), and for village-level strata. Standard errors are Huber-White heteroskedastic.

FIGURE

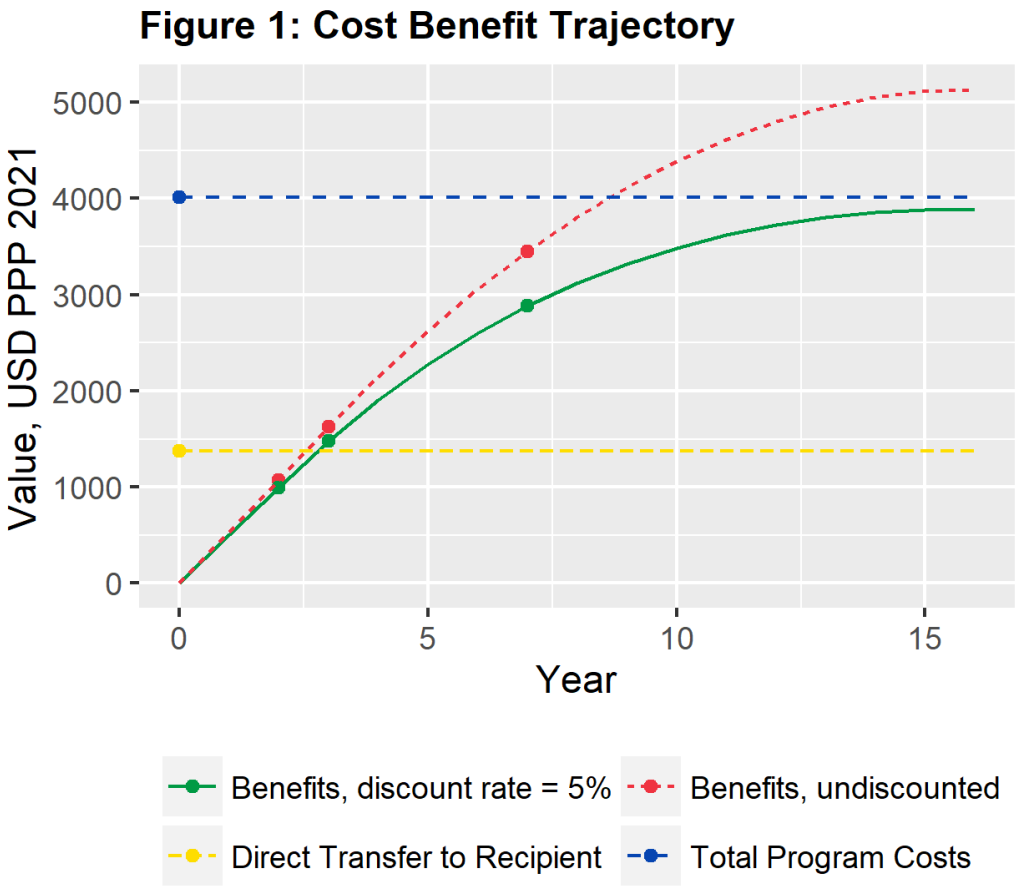

Figure 1 compares program costs and benefits. Program costs are directly reported by the implementing partner and scaled to 2021 USD in Purchasing Power Parity Terms, with a full breakdown given in Appendix Table A4. The Direct Transfer to the Recipient is the value of the asset (livestock, petty trade inputs, or beehives and bee colonies). Program benefits are calculated as monthly per capita consumption treatment effects times household size times 12. The dots represent the data points we use to estimate the trajectory (we assume a linear decay between years three and seven, and that the same trend continues until benefits reach zero). The green solid line assumes a 5% annual discount rate; the red dotted line does not apply a discount rate.

APPENDICES

Appendix A: Price Conversions

In order to ensure that any changes we observe (including among control households) reflect real differences in participant living standards, rather than changing national prices which may or may not be reflected locally, we convert all prices from Ethiopian Birr (ETB) in the year of our survey to USD 2021 in Purchasing Power Parity (PPP) terms using a basket of local prices and a multi-step procedure.

We base our basket off prices faced by sample households for food. In our consumption modules, we identify all good-unit pairs (e.g., beer bottles of cooking oil) purchased by at least five households in each of the four waves. This restriction leaves us with 22 good-unit pairs. For each good-unit pair, we calculate the median purchase price per unit (which we use as our price in the price index), and the mean amount purchased (which we use as our quantity).[1] With these values, we construct a single Fisher price index for each wave. We estimate price indices of 151, 164 and 186 for our two-, three- and seven-year surveys, respectively (with our baseline survey standardized to 100).

As a comparison, the Consumer Price Index in Ethiopia for the three follow-up years (standardizing 2010 to 100), are 156, 168, and 240, respectively. These numbers suggest that national inflation is in line with locally observed inflation in our two- and three-year surveys, but that national inflation is substantially higher by 2017.

We first use our constructed indices to convert all values from ETB in the survey year to 2010 ETB. We then convert these values to 2021 USD PPP by multiplying by 4.28, the 2010 ETB-2010 USD PPP exchange rate (per the World Bank). Finally, we multiply this value by (271.0/218.1), the ratio of the US consumer price index in 2021 over the US consumer price index in 2010.

We are unable to perform a similar transformation in India, given the limited price data collected in their surveys at the good level—their respondents were instead asked to aggregate and report the value of consumption goods at the group level (e.g., "total value of cereals and cereal products consumed in the last 30 days"). Given this, we instead take the values in BDS (reported in 2017 USD PPP) and multiply by the ratio of the US Consumer Price Index in 2021 over the US Consumer Price Index in 2017.

[1] We use medians for prices, to limit the potential influence of outliers. However, we use means rather than medians for quantities, because for most of these goods, the median amount purchased is equal to zero.

Appendix B: Cost-Benefit Simulations

We estimate the uncertainty of our cost-benefit estimates using bootstraps. In particular, we construct bootstrapped samples of our study population, re-calculate consumption benefits on the bootstrapped sub-sample, and use these calculations to re-estimate the benefit-cost ratio.

We take our full sample of 925 households ever in the sample, and in each of 10,000 simulations, randomly draw 925 households from this sample with replacement. (In each bootstrap, we sample at the household, rather than household-survey wave level. This preserves any autocorrelation that might exist between waves). We then re-estimate household benefits (equal to monthly per capita consumption average treatment effects * household size * 12) for our two-, three-, and seven-year results, and store these estimates in a matrix.

For each of these 10,000 estimates, we perform the same calculations done from our main estimates on these bootstrapped estimates to re-calculate our benefit-cost ratio. In particular, we assume that year one consumption benefits equal year two benefits, and that the benefits from years three to seven evolve linearly. In cases where the benefits decline from years three to seven, we assume that they follow a linear decline, until the point they are equal to zero.

In 19% of simulations, we find that benefits grew from years three to seven. For these simulations, we assume that benefits remain constant beyond year seven (rather than allowing these benefits to continue to grow in perpetuity). We additionally impose a cap that benefits extend for no more than 30 years.[2]

We discount all benefits with an annual rate of 5%.

Finally, for each of these 10,000 estimates, we calculate a benefit-cost ratio, by dividing the accumulated benefits over $4,011, our estimated program cost (we assume our costs are measured without error). We find that benefits exceed costs in 51.7% of our simulations; our median benefit-cost ratio is 1.02. Our 95% interval is [0.42, 2.36].

[2] This simplification is done because there are households for whom consumption estimates decline between years three and seven, albeit very slowly (our 99th percentile is an additional 363 years of benefits). Applying a perpetuity formula slightly overstates the benefits for these individuals, but there is no clear cut-off between "benefits extend forever" and "benefits ultimately reach zero." In practice, results are not very sensitive to the time horizon used in the very long run. Increasing our time horizon to 72 years increases the share for whom benefits exceed costs from 51.72% to 51.90%. Extending the benefits beyond year 72 is not pivotal for whether any additional households have benefits in excess of program costs.

APPENDIX TABLES

Appendix Table A1: Baseline Balance Tests

	(1)	(2)	(3)	(4)	(5)
Panel A. Household-Level Outcomes					
	Asset Ownership	Per Capita Consumption	Food Security Index	Income and Revenues	Financial Inclusion
Treatment	0.098	-0.059	0.026	0.040	-0.085
	(0.063)	(0.063)	(0.064)	(0.064)	(0.065)
Observations	925	925	925	925	925
Control mean	-0.045	0.028	-0.010	-0.019	0.042
Panel B. Adult-Level Outcomes					
	Physical Health Index	Mental Health Index	Time Working	Political Involvement Index	Women's Empowerment Index
Treatment (ITT): Three Year	-0.081	0.026	-0.051	0.003	-0.099
	(0.059)	(0.059)	(0.058)	(0.055)	(0.069)
Observations	1305	1304	1305	1305	773
Control Mean	0.042	0.042	0.042	0.042	0.042

Appendix Table A1 tests for balance on baseline outcomes between treatment and control. Each regression controls for village-level strata. Standard errors are Huber-White heteroskedastic. Each outcome is standardized to have mean 1, standard deviation 0.

Appendix Table A2: Program Timeline

Activity	Date
Baseline Survey	April 2010
Public Lottery	May 2010
Asset Transfer	June-August 2010, December 2010*
Training and Coaching	June 2010-May 2012
Endline Survey 1	July 2012
Endline Survey 2	July-August 2013
Endline Survey 3	September 2017

Initital transfers took place from June-August. Due to concerns livestock purchases would drive up the price of livestock (and reduce the number of livestock that could be transferred), households receiving sheep and goats received half their livestock in June-August, and the other half in December

Appendix Table A3: Attrition

	(1) 2 Years	(2) 3 Years	(3) 7 Years
Panel A: Attrition by Treatment Status			
Treatment	0.000	-0.006	-0.008
	(0.01)	(0.01)	(0.01)
Control Mean	0.0107	0.0214	0.0428
Panel B: Correlates of Attrition			
Asset Ownership Index	0.000	0.000	0.007
	(0.004)	(0.005)	(0.008)
Per Capita Consumption Index	0.000	-0.005	-0.007
	(0.004)	(0.005)	(0.008)
Food Security Index	-0.001	0.002	0.006
	(0.004)	(0.005)	(0.007)
Income and Revenues Index	-0.001	-0.002	0.001
	(0.004)	(0.005)	(0.008)
Financial Inclusion Index	0.001	0.002	0.010
	(0.004)	(0.005)	(0.007)
Physical Health Index, HH-Level Average	0.00611*	0.002	0.000
	(0.004)	(0.005)	(0.008)
Mental Health Index, HH-Level Average	0.003	0.002	0.001
	(0.003)	(0.004)	(0.007)
Productive Time Use Index, HH-Level Average	-0.004	-0.006	-0.011
	(0.003)	(0.005)	(0.007)
Political Involvement Index, HH-Level Average	0.001	0.003	-0.005
	(0.004)	(0.005)	(0.007)
Women's Empowerment Index	-0.002	-0.001	0.000
	(0.003)	(0.004)	(0.006)
Panel C: Test for Differential Composition of Attriters by Treatment			
Joint F-Test: Treatment and Indices Interacted with Treatment Status	0.350	0.353	0.790
p-value	0.984	0.983	0.671

Panel A reports regression results of whether or not an individual attrited from the sample on treatment status, with attrition as the dependent variable. Panel B regresses attrition on the baseline values of our indexed family outcomes. Panel C reports the joint F-Test from a regression of attrition on the correlates in Panel B interacted with treatment. In all cases, we include village-level strata indicator variables; standard errors are Huber-White heteroskedastic.

Appendix Table A4: All Indexed Family Outcomes, Ethiopia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Asset Ownership	Per Capita Consumption	Food Security Index	Income and Revenues Index	Financial Inclusion Index	Physical Health Index	Mental Health Index	Time Working	Political Involvement Index	Women's Empowerment Index
Treatment (ITT): Two Year	0.950***	0.239***	0.116**	1.411***	1.853***	-0.0158	0.0527	0.278***	0.0969*	-0.0155
	(0.09)	(0.07)	(0.05)	(0.16)	(0.12)	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)
q-value	0.001	0.001	0.022	0.001	0.001	0.101	0.049	0.108	0.112	0.486
Control mean	0.374	0.00	0.706	0.00	0.00	-0.0749	0.00	0.508	0.442	0.188
Observations	915	915	915	915	915	1,307	1,307	834	1,307	761
Treatment (ITT): Three Year	0.947***	0.250***	0.145***	0.411***	0.779***	0.0499	-0.0384	0.224***	0.115*	-0.0266
	(0.09)	(0.05)	(0.05)	(0.12)	(0.13)	(0.04)	(0.06)	(0.06)	(0.06)	(0.05)
q-value	0.001	0.001	0.003	0.001	0.001	0.117	0.050	0.104	0.115	0.431
Control mean	0.519	-0.0948	0.742	0.512	0.0473	0.0112	0.00	0.604	0.235	0.241
Observations	908	908	908	915	908	1,267	1,265	825	1,265	726
Treatment (ITT): Seven Year	0.430***	0.172**	0.0351	0.244	0.301	-0.116	-0.0872	0.0805	0.00662	-0.0544
	(0.11)	(0.08)	(0.04)	(0.24)	(0.25)	(0.08)	(0.06)	(0.05)	(0.06)	(0.06)
q-value	0.001	0.196	0.464	0.464	0.449	0.109	0.046	0.191	0.074	0.128
Control mean	1.159	0.487	1.024	1.247	0.736	-0.0525	0.00	0.463	0.214	0.572
Observations	889	889	889	889	889	1,228	1,228	1,228	1,228	723
Baseline Mean	0.00	-	0.00	-	-	0.01	-	0.00	0.00	-0.01

Appendix Table A4 reports results from intention-to-treat effects of the graduation program on indexed family outcomes. For asset ownership, consumption and time use, outcomes are aggregated and then rescaled; other outcomes are a z-score index of outcomes variables in the given category. Outcomes are either (a) standardized so the baseline has mean zero and standard deviation one (done whenever possible), or (b) standardized so the endline 1 control group has mean zero and standard deviation one (done in cases where we do not have baseline data, or the baseline components differed from the components in subsequent waves). Reported q-values are sharpened using the false discovery rate procedure detailed in Anderson (2008). The full list of variables used to construct each index is reported in Appendix A. Each regression controls for baseline values, and for village-level strata. Standard errors are Huber-White heteroskedastic for household-level outcomes; adult-level outcomes are clustered at the household level.

Appendix Table A5: Cost-benefit analysis

Panel A: Program Costs per Household, USD PPP 2021		
(1)	Direct Transfer Costs	1371
	Asset Cost	1371
	Food stipend	0
	Total Supervision Costs	2122
	Salaries of Implementing Organization Staff	387
	Materials	37
	Training	949
	Travel Costs	194
	Other Supervision Expenses	554
	Total Direct Costs	3493
	Start-up expenses	48
	Indirect Costs	470
(2)	Total Costs, calculated as if all incurred immediately at beginning of Year 0	4011
Panel B: Benefits per Household, USD PPP, All Values Deflated to Baseline at 5% annual social discount rate		
(3)	Year 1 Annual Consumption ITT, assuming treatment effect equal to Year 2	509
(4)	Year 2 Annual Consumption ITT Treatment Effect	484
(5)	Year 3 Annual Consumption ITT Treatment Effect	484
(6)	Year 7 Annual Consumption ITT Treatment Effect	280
	Estimated Benefits, Years 4-6, assuming linear decay from Year 3 to 7	1125
(7)	Projected Future Benefits, Years 8-16, assuming linear decay from Year 3 to Year 7 continues	1012
(8)	Total Estimated and Projected Consumption Benefits, (3) + (4) + (5) + (6) + (7)	3894
Panel C: Benefit-Cost Ratio		
	Total Consumption Benefits divided by Costs, (8) / (2)	0.971
	Total Consumption Benefits divided by Direct Transfer to Recipients, (8) / (1)	2.839
	Consumption Benefits Realized by Year 7, Divided by Costs, ((8) - (7)) / (1)	0.718
	Total Consumption Benefits divided by Costs, if not discounting future	1.233
	Share of bootstraps in which benefits exceed costs	0.519
	Median benefit-cost ratio from bootstrapped estimates	1.024
	95% Confidence Interval for benefit-cost ratio from bootstrapped estimates	[0.42, 2.36]

Appendix Table A5 presents cost-benefit estimates. Costs are reported by the implementing partner, and converted to 2021 USD in PPP terms. Benefits are calculated as equal to the sum of accumulated and projected future consumption benefits. We calculate the benefits in years 2, 3, and 7 from our consumption modules (and scaled to annual values). We assume that the decay in consumption benefits from Year 3 to 7 is linear, and that this decay continues linearly until the benefits reach 0, by Year 16.

Appendix Table A6: Asset Ownership

	(1)	(2)	(3)
	Total Asset Value, USD PPP 2021	Productive Asset Value, USD 2021	Durable Good Value, USD 2021
Treatment (ITT): Two Year	1,162***	931.5***	61.77***
	(111.60)	(90.95)	(20.02)
Control mean	1659	1239	206.7
Observations	915	915	910
Treatment (ITT): Three Year	1,158***	913.7***	72.51***
	(109.40)	(88.37)	(17.76)
Control mean	1835	1364	205.9
Observations	908	908	908
Treatment (ITT): Seven Year	525.3***	448.4***	84.11**
	(133.30)	(122.40)	(39.36)
Control mean	2618	2193	424.2
Observations	889	889	889
Baseline Mean	1200.97	1156.83	44.14

Appendix Table A6 reports results from intention-to-treat effects of the graduation program on asset ownership (and its component parts), the indexed version of which is reported in Table 1. Each regression controls for baseline values (which we construct using relative prices in subsequent waves--we observe baseline quantities of the same asset types, but not baseline values.). We additionally control for village-level strata. Standard errors are Huber-White heteroskedastic.

Appendix Table A7: Food Security

	(1)	(2)	(3)	(4)	(5)
	Everyone gets enough food everyday	No Adults skipped meals	No HH member went a whole day without food	No children skipped meals	Everyone eats at least two meals everyday
Treatment (ITT): Two Year	0.0386	0.0628**	0.0293*	0.0447**	0.0178
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Control mean	0.64	0.656	0.913	0.845	0.91
Observations	914	910	910	910	909
Treatment (ITT): Three Year	0.0766**	0.0513*	0.0411***	0.0468**	0.025
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Control mean	0.64	0.687	0.921	0.833	0.923
Observations	907	907	906	904	906
Treatment (ITT): Seven Year	0.0234	0.0211	-0.00719	0.0113	0.0118
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Control mean	0.826	0.850	0.987	0.926	0.908
Observations	889	889	889	802	889
Baseline Mean	0.34	0.37	0.75	0.49	0.83

Appendix Table A7 reports results from intention-to-treat effects of the graduation program on food security. A subset of these variables is reported in Table 2, replicating the procedure of Banerjee et al. 2021. Each of these variables is used to construct the food security index in Table 1. Each regression controls for baseline values (where available), and for village-level strata. Standard errors are Huber-White heteroskedastic.

Appendix Table A8: Financial Inclusion

	(1)	(2)	(3)	(4)
	Amount borrowed, last 12 months, formal sources, USD PPP 2021	Amount borrowed, last 12 months, informal sources, USD PPP 2021	Savings balance, USD PPP 2021	Savings deposits, last 3 months, USD PPP 2021
Treatment (ITT): Two Year	37.30**	19.56	849.7***	43.58***
	(14.89)	(16.53)	(38.55)	(4.44)
Control mean	22.34	117.3	91.13	10.06
Observations	915	915	915	915
Treatment (ITT): Three Year	41.99**	28.87**	315.9***	9.472*
	(19.66)	(11.29)	(41.88)	(4.95)
Control mean	34.61	134.8	84.87	7.551
Observations	908	908	908	908
Treatment (ITT): Seven Year	-8.961	26.65	75.94	17.07
	(36.52)	(45.88)	(55.29)	(14.41)
Control mean	102.2	55.62	244.4	43.91
Observations	889	889	889	889
Baseline Mean	42.32	4.70	-	-

Appendix Table A8 reports results from intention-to-treat effects of the graduation program on financial inclusion. Each of these variables is used to construct the financial inclusion index in Table 1. Each regression controls for baseline values (where available), and for village-level strata. Standard errors are Huber-White heteroskedastic.

Appendix Table A9: Productive Time Use

	(1)	(2)	(3)	(4)	(5)
	Minutes working per day on average, last 48 hours	Minutes working in agriculture on average, last 48 hours	Minutes working with livestock on average, last 48 hours	Minutes working on non-farm enterprise on average, last	Minutes working in wage labor on average, last 48 hours
Treatment (ITT): Two Year	58.54***	21.30**	37.77***	6.438*	-6.142
	(11.97)	(9.34)	(7.58)	(3.54)	(6.35)
Control mean	224.50	96.39	80.30	2.86	44.96
Observations	834	834	834	834	834
Treatment (ITT): Three Year	47.07***	22.57**	30.93***	4.41	-9.801
	(12.43)	(10.55)	(7.88)	(3.03)	(6.78)
Control mean	244.80	125.00	75.73	3.59	40.44
Observations	825	825	825	825	825
Treatment (ITT): Seven Year	16.92	8.22	6.516	3.403	0.176
	(10.40)	(9.04)	(6.90)	(2.65)	(4.43)
Control mean	215.10	102.60	88.07	3.96	20.41
Observations	1,228	1,228	1,228	1,228	1,228
Baseline Mean	118.14	18.50	39.62	2.67	57.46

Appendix Table A9 reports results from intention-to-treat effects of the graduation program on productive time use. The aggregate (in column 1) is reported in Table 1 (rescaled to have baseline mean 0 and standard deviation 1). Each regression controls for baseline values, and for village-level strata. Standard errors are clustered at the household level.

Appendix Table A10: Physical and Mental Health

	Physical Health Variables			*Mental Health Variables*		
	(1)	(2)	(3)	(4)	(5)	(6)
	No days of work missed due to poor physical health	Mean Score, Activites of Daily Living (0/1)	Perception of Physical Health (1/5)	Overall Satisfaction with Life (1/5)	No extended period of time with worry	Stress Index
Treatment (ITT): Two Year	0.00432	-0.00651	-0.0293	0.146**	0.00381	-0.00667
	(0.01)	(0.02)	(0.07)	(0.07)	(0.02)	(0.05)
Control mean	0.949	0.861	3.62	3.516	0.86	0.00
Observations	1,301	1,292	1,292	1,307	1,307	1,307
Treatment (ITT): Three Year	0.00538	0.0283**	-0.0282	0.083	-0.0246	-0.0585
	(0.01)	(0.01)	(0.06)	(0.06)	(0.02)	(0.06)
Control mean	0.959	0.879	3.731	3.55	0.939	0.00
Observations	1,263	1,256	1,256	1,265	1,264	1,265
Treatment (ITT): Seven Year	0.00193	-0.0183	-0.229	-0.168	-0.00925	-0.0663
	(0.01)	(0.02)	(0.14)	(0.14)	(0.02)	(0.07)
Control mean	0.0455	0.848	7.087	7.307	0.896	0.00
Observations	1,228	1,227	1,228	1,228	1,227	1,228
Baseline Mean	0.907	0.874	3.938	3.518	0.854	0.009

Appendix Table A10 reports results from intention-to-treat effects of the graduation program on physical and mental health outcomes. The physical health index in Table 1 is comprised of the variables in columns 1-3, while the mental health index is comprised of the variables in columns 4-6. Each regression controls for baseline values, and for village-level strata. Standard errors are clustered at the household level.

Appendix Table A11: Political and Women's Empowerment

	Political Involvement			*Women's Decision-Making*				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Individual has attended meeting with local leader or politician	Individual has asked a question of local leader or politician at meeting	Individual is member of political party	Women has major say in food-related spending decisions in household	Women has major say in education-related spending decisions in household	Women has major say in healthcare-related spending decisions in household	Women has major say in home improvement spending decisions in household	Women has major say in household management decisions in household
Treatment (ITT): Two Year	0.0466*	-0.0092	0.0605**	-0.0451	-0.00685	-0.0127	0.025	-0.02
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Control mean	0.613	0.296	0.349	0.735	0.491	0.556	0.406	0.402
Observations	1,307	1,307	1,306	758	742	757	755	745
Treatment (ITT): Three Year	0.0474*	0.00401	0.0609**	0.00975	-0.00946	-0.0148	-0.0234	-0.0301
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Control mean	0.522	0.216	0.328	0.78	0.492	0.563	0.447	0.393
Observations	1,262	1,265	1,265	711	717	725	722	655
Treatment (ITT): Seven Year	0.0119	-0.00254	-0.00122	0.00263	-0.0252	-0.0253	-0.0464	-0.0407
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Control mean	0.51	0.192	0.349	0.836	0.652	0.661	0.707	0.615
Observations	1,227	1,227	1,228	723	723	723	723	723
Baseline Mean	0.350	0.184	0.303	0.473	0.411	0.446	0.409	0.383

Appendix Table A11 reports results from intention-to-treat effects of the graduation program on physical and mental health outcomes. The political empowerment index in Appendix Table A3 is comprised of the variables in columns 1-3, while the women's empowerment index is comprised of the variables in columns 4-8. Each regression controls for baseline values, and for village-level strata. Standard errors are clustered at the household level.